Good + Bad = ? Developmental Differences in Balancing Gains and Losses in Value-Based Decisions From Memory

Horn, Sebastian S ; Mata, Rui ; Pachur, Thorsten

Abstract: Value-based decisions often involve comparisons between benefits and costs that must be retrieved from memory. To investigate the development of value-based decisions, 9- to 10-year olds (N = 30), 11- to 12-year olds (N = 30), and young adults (N = 30) first learned to associate gain and loss magnitudes with symbols. In a subsequent decision task, participants rapidly evaluated objects that consisted of combinations of these symbols. All age groups achieved high decision performance and were sensitive to gain–loss magnitudes, suggesting that required core cognitive abilities are developed early. A cognitive-modeling analysis of performance revealed that children were less efficient in object evaluation (drift rate) and had longer nondecision times than adults. Developmental differences, which emerged particularly for objects of high positive net value, were linked to mnemonic and numerical abilities.

DOI: https://doi.org/10.1111/cdev.13208

Originally published at:
Horn, Sebastian S; Mata, Rui; Pachur, Thorsten (2020). Good + Bad = ? Developmental Differences in Balancing Gains and Losses in Value-Based Decisions From Memory. Child Development, 91(2):417-438.
DOI: https://doi.org/10.1111/cdev.13208

Good + Bad = ? Developmental Differences in Balancing Gains and Losses in Decisions from Memory

Sebastian S. Horn

Max Planck Institute for Human Development and University of Zurich

Rui Mata

University of Basel

Thorsten Pachur

Max Planck Institute for Human Development

Author Note

Sebastian S. Horn, Center for Adaptive Rationality (ARC), Max Planck Institute for Human Development Berlin, Germany, and University of Zurich, Switzerland. Rui Mata, Center for Cognitive and Decision Sciences, Department of Psychology, University of Basel, Switzerland. Thorsten Pachur, Center for Adaptive Rationality (ARC), Max Planck Institute for Human Development, Berlin, Germany.

This research was supported by a fellowship from the Max Planck Society to Sebastian Horn. We thank Ksenia Appelganc, Lilla Horvath, Eva Woldrich, Veronika Zilker for help with data collection and Guido Biele for providing us with stimulus materials. We also thank Valerie Reyna, Roger Ratcliff, Hilde Huizenga, Bernd Figner, Eldad Yechiam, and the members of the Center for Adaptive Rationality (ARC), for helpful comments on parts of this project.

Correspondence concerning this article should be addressed to Sebastian Horn, who is now at the University of Zurich, Department of Psychology, Binzmuehlestrasse 14 (Box 11), 8050 Zurich, Switzerland. E-mail: horn@psychologie.uzh.ch

Abstract

Value-based decisions often involve comparisons between benefits and costs that must be retrieved from memory. To investigate the development of value-based decisions, 9-to-10-year-olds (*N*=30), 11-to-12-year-olds (*N*=30), and young adults (*N*=30) first learned to associate gain and loss magnitudes with symbols. In a subsequent decision task, participants rapidly evaluated objects that consisted of combinations of these symbols. All age groups achieved high decision performance and were sensitive to gain–loss magnitudes, suggesting that required core cognitive abilities are developed early. A cognitive-modeling analysis of performance revealed that children were less efficient in object evaluation (drift rate) and had longer nondecision times than adults. Developmental differences, which emerged particularly for objects of high positive net value, were linked to mnemonic and numerical abilities.

Keywords: cognitive development, decision making, gain–loss valuation, diffusion model, response times

Supplemental materials for this article are available online.

Good + Bad = ? Developmental Differences

in Balancing Gains and Losses in Decisions from Memory

Decision makers of all ages encounter situations that have both beneficial and costly consequences. Decisions about what we like or prefer (value-based decisions) thus often involve a consideration among competing attributes. Relevant information about benefits and costs, however, is rarely directly provided in the environment but must be retrieved from memory. Preference construction from memory can explain a wide variety of behavioral phenomena (Reyna, Lloyd, & Brainerd, 2003; Stewart, Chater, & Brown, 2006; Weber & Johnson, 2006). For example, children may face conflict of approach versus avoidance because objects or situations are associated with both positive and negative valence based on past experience (Lewin, 1935). Constructing subjective value from memory and comparing attributes to evaluate utility involves, in particular, associative-memory processes (Shadlen & Shohamy, 2016). Neuroeconomic research has demonstrated the relevance of medial temporal brain regions for these processes (Enkavi et al., 2017; Glimcher & Fehr, 2014; Gluth, Sommer, Rieskamp, & Büchel, 2015). Based on these findings and on the premise that brain structures relevant for associative remembering are developed early in the lifespan (e.g., Shing et al., 2010), we investigate the development of rapid value-based decisions based on memory representations.

In this study, we test how well children and early adolescents, relative to adults, perform value-based decisions that involve elementary gain–loss comparisons. In younger adults, decisions involving both gains and losses tap into difference-based representations of subjective value and are more accurate and rapid the larger the net value (Basten, Biele, Heekeren, & Fiebach, 2010). How such regularities develop and how they are associated with developmental cognitive change, however, is largely unknown. We analyze the underlying processes of value-based decisions with a drift-diffusion model, assuming that preferences can be constructed through sequential sampling from memory (Johnson & Ratcliff, 2013; Konovalov & Krajbich, 2017; Rangel, Camerer, & Montague, 2008). This modeling approach disentangles psychologically important components of decision making and allows us to account for developmental differences in all aspects of the behavioral data, including response time (RT). Moreover, we explore the mnemonic and other cognitive abilities that these decisions recruit.

In what follows, we first discuss the relevance of associative memory in value-based decision making. We then review research on adults' and children's gain–loss valuation and identify open questions regarding the cognitive abilities that may underlie developmental differences. Finally, we describe the paradigm and modeling approach that we use to investigate the development of value-based decisions from memory and report a study with fourth-grade children, sixth-grade early adolescents, and young adults.

The Development of Associative Memory

To evaluate objects, their positive and negative attributes are often derived by probing memory for past experience. Value-based decisions are therefore closely interwoven with memory and learning processes (see Shadlen & Shohamy, 2016). Specifically, value-based decisions likely involve *associative memory*, which refers to processes relevant (a) for retrieving the links between objects, persons, or situations, and their experienced values (Gluth et al., 2015) and (b) for binding these features into a compound mnemonic representation. What differences between children and adults can be expected? Developmental research indicates that associative memory functioning (which largely involves the medial temporal lobe system) is relatively mature by middle childhood (see Shing et al., 2010, for an overview). Four-year-olds already show a remarkable ability to bind information in memory (Sluzenski, Newcombe, & Kovacs, 2006). To the extent that value-based decisions tap into associative memory processes, one may thus expect younger children already to perform such decisions relatively well.

Relatedly, developmental theories about the influence of different memory representations on decision making have emphasized a developmental trend from childhood to adulthood from predominant reliance on quantitative to qualitative evaluation. Specifically, fuzzy-trace theory (e.g., Reyna, 2012; Reyna et al., 2003) suggests that children—within the limits of their computational capabilities—tend to rely relatively more on verbatim than on gist memory representations (which support exact quantitative vs. meaning-based qualitative evaluation, respectively); hence, even though accuracy of both types of memories increases from childhood to adulthood, it is conceivable from this perspective that even younger children can perform relatively accurately if decisions require matching perceptual information with associated values.

The Development of Gain–Loss Comparison in Decision Making

Motivational Sensitivity to Gains and Losses

The valuation of positive and negative attributes has received considerable attention in research on judgment and decision making with adults (e.g., Bechara, Damasio, Damasio, & Anderson, 1994; Gächter, Johnson, & Herrmann, 2007; Slovic & Lichtenstein, 1968; Tversky & Kahneman, 1992). One key notion in this research is that people often perceive and treat gains and losses differently. For instance, losing $100 may hurt more than winning $100 feels good. The concept of *loss aversion* implies that losses have a stronger impact on decisions than equivalent gains (Tversky & Kahneman, 1991). Adults have shown behavior consistent with loss aversion in various studies (e.g., people typically avoid symmetric bets in which an amount of *x* or –*x*, respectively, can be gained or lost with equal probability; Tom, Fox, Trepel, & Poldrack, 2007). However, recent research suggests that loss aversion does not operate in all situations involving losses: for example, when gains and losses occur simultaneously (or sequentially within a short time), when only small-to-moderate amounts of money are at stake (Harinck, Van Dijk, Van Beest, & Mersmann, 2007), or when outcomes are learned from experience, there is sometimes no loss aversion in adults (Erev, Ert, & Yechiam, 2008; Yechiam & Hochman, 2013, for an overview).

How does the processing of gains and losses in decision making develop? Little is known about developmental differences in riskless gain–loss decisions from memory. In contrast, a wealth of research has examined the development of decision making involving trade-offs between gains and losses under risk (where the outcomes occur with some clearly specified probability; Boyer, 2006; Defoe, Dubas, Figner, & van Aken, 2015; Reyna, 2012). Moreover, developmental research has examined decisions in which information about probabilities is incomplete or ambiguous (e.g., Tymula et al., 2012) or in which both outcomes and their frequencies are uncertain and must be explored through experience (Rosenbaum, Venkatraman, Steinberg, & Chein, 2017; van den Bos & Hertwig, 2017). Further developmental studies have addressed how punishment and reward influences children's decisions in reinforcement learning (Costantini & Hoving, 1973; Hämmerer, Li, Müller, & Lindenberger, 2010; van Duijvenvoorde, Zanolie, Rombouts, Raijmakers, & Crone, 2008). One overarching finding emerging from this research is that motivational sensitivity towards outcomes (e.g., tokens or monetary value) changes fundamentally from childhood to adulthood on both the behavioral and brain level (Beitz, Salthouse, & Davis, 2014; Hämmerer et al., 2010; van Duijvenvoorde et al., 2008, 2015; Weller, Levin, & Denburg, 2011).

On the one side, there is evidence that younger children already choose as if losses had a relatively stronger impact than gains. Like adults, for example, preschool children make riskier choices to avoid losses than to achieve gains in reflection problems (e.g., Reyna, 2012). Greater impact of negative than of positive feedback on children's decisions has also been observed in reinforcement learning (Costantini & Hoving, 1973; Hämmerer et al., 2010). Moreover, children as young as 0 years show endowment effects in trading behavior, presumably reflecting a stronger psychological impact of losing an endowed good over gaining another (Harbaugh, Krause, & Vesterlund, 2001).

On the other side—and at odds with key notions in decision research with adults (Tversky & Kahneman, 1991)—gains (rather than losses) appear to have a particularly strong motivational influence on decisions during specific developmental phases. Neurodevelopmental models suggest that this might be due to an imbalance between reward-related processing and cognitive control, which leads to strong responsivity to reward during adolescence (Somerville, Jones, & Casey, 2010; Steinberg, 2008). For example, Cauffman et al. (2011) examined approach behavior (operationalized as tendency to play increasingly from advantageous decks) and avoidance behavior (tendency to refrain from playing from disadvantageous decks) during the Iowa Gambling Task, which requires choices among options (card decks) with both gain and loss outcomes. Cauffman et al. reported that approach toward potential gains followed a curvilinear pattern across development, with sensitivity to positive feedback (to achieve gains) peaking during the adolescent years. Further research suggests that the attractiveness of gains in adolescence could be particularly prominent when decision problems involve active exploration (sampling of outcomes through experience; Rosenbaum et al., 2017) in heightened affective contexts (Figner, Mackinlay, Wilkening, & Weber, 2009).

Taken together, the findings about specific positivity or negativity effects in gain–loss decisions appear complex: Several studies indicate that loss aversion emerges relatively early in childhood in risky (Reyna, 2012) and riskless decisions (Harbaugh et al., 2001) and that motivational sensitivity to losses may remain relatively stable across the lifespan (Beitz et al., 2014; Weller et al., 2011). There is also evidence, however, that gains have a stronger impact than losses on value-based decisions during adolescence, possibly depending on how information about outcomes is acquired (Figner et al., 2009; Rosenbaum et al., 2017). For the present investigation of value-based decisions, this implies that losses may not necessarily loom larger than gains in tasks in which gains and losses are concurrently remembered and compared (Yechiam & Hochman, 2013). Moreover, it seems important to distinguish motivational sensitivity to gains (reward) and losses (punishment) from the sensitivity to numerical quantities, which may follow different developmental trajectories. We therefore turn to the role of numerical and other cognitive abilities that could also be important for understanding developmental differences in value-based decision tasks.

Numerical Abilities

Decisions about values typically involve the evaluation of quantitative information and likely tap into abilities related to numerical processing and magnitude representation. If age-tailored materials are used under simple presentation conditions, preschool children (around age five) already show the basic competence of evaluating both outcomes and their probabilities (e.g., Schlottmann & Wilkening, 2011). Moreover, within their capabilities, children appear to focus relatively more on quantitative details than on interpretation of the meaning of outcomes in decision making (Reyna & Brainerd, 1994).

However, numerical abilities have rarely been assessed in previous developmental investigations of value-based decisions, in which children were exposed to rewards of various magnitudes (e.g., Schlottmann & Wilkening, 2011). For adults, it has been shown that their perception of symbolic numerals (e.g., “40”) influences their subjective valuation (e.g., of a monetary value of $40): People with more exact symbolic-number mappings have more linear value functions (Schley & Peters, 2013). Relatedly, a wealth of developmental research indicates that the precision of basic numerical intuitions (number sense) continues to increase monotonically throughout the school-age years, with highest levels of acuity attained surprisingly late in adulthood (e.g., Halberda, Ly, Wilmer, Naiman, Germine, 2012). That is, humans acquire increasingly precise representations of nonsymbolic numerical magnitudes and their relation to symbolic numerals (e.g., Arabic numerals, “1”, “5”, etc.) throughout development. Moreover, the range of accurately represented natural numbers increases continuously during childhood (e.g., 8- to 10-year-olds gain experience with the 0–1000 range; 10- to 12-year-olds with the 0–10,000 range). For an overview, see Siegler (2016).

Given the potential relevance of these numerical abilities in gain–loss comparisons, we explored to what extent they might mediate age differences in value-based decisions. In fact, children may be less sensitive than adults to differences between magnitudes because they have a less exact (more compressed) mental representation of values. Such developmental differences could be particularly pronounced for higher magnitudes, with which children have less experience than adults. Thompson, Ratcliff, and McKoon (2016) found in modeling analyses that school-age children were less efficient than adults in extracting information from both symbolic and nonsymbolic numeric stimuli; these age differences were more pronounced when numeric comparisons involved more extreme (higher or lower) values in terms of their numerical distance from an internal reference point (numerical distance effect). To the extent that value-based decisions also recruit such basic numerical skills, children might have more difficulties accurately integrating gains and losses and their sensitivity to the net values of different objects could be less pronounced than in adults.

Fluid Cognitive Abilities

A wealth of developmental research has shown that elementary aspects of information processing strongly improve as children mature (see Fry & Hale, 2000). Hence, intellectual performance on a wide variety of decision tasks—including value-based choice—may depend on variables such as speed, working memory, and fluid intelligence. To evaluate the proposed influence of mnemonic access and numerical abilities on value-based choice above and beyond such elementary capabilities of information processing, we also collected measures of participants' fluid abilities to simultaneously account for their potential impact on the speed and accuracy of decision making.

Overview of the Current Study

To investigate the development of value-based decisions from memory, we employed a child-friendly, modified version of a decision paradigm that separates the learning of outcomes from decision making (Basten et al., 2010). We focused on the ages between 9-12 years because this has been shown to be an important range for developmental changes in the evaluation of positive and negative information (e.g., van Duijvenvoorde et al., 2008) and numerical competencies (Thompson et al., 2016) that we expected to contribute to the balancing of gains and losses. In our study, school-aged children, early adolescents, and young adults first experienced associations between symbols (colors or shapes) and their value (different gain or loss magnitudes). In a subsequent decision task, participants then saw individual objects (each representing a combination of two symbols) and were asked to accept or reject these objects (that yielded overall net gain or net loss).

Extant developmental studies involving gain or loss outcomes have focused almost exclusively on choice patterns (for overviews, see Boyer, 2006; Defoe et al., 2015; Reyna & Brainerd, 1994; Rosenbaum et al., 2017; Schlottmann & Wilkening, 2011; van Duijvenvoorde, Jansen, & Huizenga, 2015). Here, we applied diffusion modeling that also takes into account the time dynamics of decisions. Compared to modeling approaches that focus on choice alone, diffusion modeling provides an improved and more comprehensive approach to detect and understand developmental differences in value-based decision making. This is because response time can provide important information when choices do not differentiate and because there are substantial changes in speed of processing from childhood to adolescence (Fry & Hale, 2000; Kail, 1991) that can have multiple and different reasons (Thompson et al., 2016).

Drift-Diffusion Modeling

Diffusion models (Ratcliff, 1978) have been applied in many investigations of rapid perceptual (Ratcliff & McKoon, 2008), memory-based (Spaniol, Voss, & Grady, 2008), and value-based decisions (Basten et al., 2010; Gluth et al., 2015). A core assumption is that decisions result from continuous sampling of evidence over time, moving from a starting point z until one of two decision boundaries is passed (Figure 1). Sampling processes are noisy thus generating distributions of decision latencies across trials. The diffusion model disentangles psychologically meaningful parameters of decision making whose interpretation has been validated in selective-influence studies (Ratcliff & McKoon, 2008; Voss, Rothermund, & Voss, 2004). In the current analysis of value-based decisions, the following model parameters are of particular interest:

The *drift rate v* quantifies the speed of information uptake. Lower absolute values of v imply slower and less accurate decisions. In the present task, drift rate is positive (negative) if relatively more evidence is accumulated in favor of acceptance (rejection) of an object, associated with the upper (lower) decision boundary in the model. Hence, drift rate quantifies the quality of match between a presented object and memory (Ratcliff, 1978) and reflects the net evidence resulting from integration of that information from memory (retrieved positive and negative attributes; Basten et al., 2010). With decreasing net value (distance) between two attributes of an object, the drift rate decreases, people make slower decisions, and they approach indifference between response options (Busemeyer & Diederich, 2002; Johnson & Ratcliff, 2014; Konovalov & Krajbich, 2017). For example, value-based decisions about an object $O_1 = [+100, -20]$, associated with +100 gain and −20 loss units, are easier than decisions about an object $O_2 = [+100, -60]$ of +100 gain and −60 loss units (points, tokens, or monetary value). Moreover, decisions about objects with the same absolute net values are equally difficult if gain and loss attributes are remembered and integrated equally well: For example, decisions about an object $O_2 = [+100, -60]$ and about an object $O_3 = [+60, -100]$ should be equally difficult under this assumption. Therefore, the comparison of drift rate for objects of equal net value—but of different sign—is used to assess people's gain–loss attitudes during information accumulation and indicates whether this information is sampled differently for gains and losses (e.g., Clay, Clithero, Harris, & Reed, 2017). Consequently, we quantify *stimulus-evaluation bias* by calculating the sum of the drift rates for gain and loss objects of equal absolute value: $v_{\text{bias}} = v_{\text{gain}} + v_{\text{loss}}$ (see also Spaniol et al., 2008). If v_{bias} is close to zero, this implies that gains and losses are evaluated similarly; if v_{bias} is positive (negative), this implies a relatively stronger impact of gains (losses) during decision making. For example, people who score low on indices of loss aversion (e.g., in standard lottery tasks) also give more weight to gain- than to loss-related attributes in rapid decision tasks and accumulate gain-related information at a relatively higher rate (Clay et al., 2017).

The *boundary separation* (parameter a) quantifies the amount of evidence required until a decision is made (reflecting a decision maker's speed–accuracy setting or cautiousness; by decreasing a, accuracy is traded for speed, and vice versa). One assumption is that decision makers set their stopping rule (boundary) prior to the sampling process. In the current task, for example, people might set wider boundaries if they require more evidence from memory to reduce uncertainty before they make a decision. The relative position of the starting point between these boundaries (z/a) quantifies the amount of evidence required for each decision. Therefore, systematic deviations from $z/a = .50$ (equidistance between boundaries) would indicate an a priori *response bias* (favoring a particular response over the other): Values of $z/a >$ .50 would imply a general tendency to accept objects whereas values of $z/a <$ .50 would imply a tendency to reject objects. For example, people with a stronger tendency to avoid losses (or to maintain their status quo) might require asymmetrically more evidence to accept than to reject an object in the current task.

The *nondecision time* T_{er} quantifies the latency of processes before and after the actual decision phase: The initial perceptual encoding of presented stimuli (transformation into decision-relevant information) and motor response execution following the decision are all combined into parameter T_{er}. Changes in T_{er} shift the entire RT distributions without affecting accuracy. Finally, by assuming variability across trials in drift rates (with standard deviation η), in starting points (with range s_z), and in nondecision times (with range s_t), the model can account for systematic differences in the RT distributions of correct and error decisions.

Research Questions

Economic and decision research with adults has shown the importance of associative remembering in value-based decisions (e.g., Enkavi et al., 2017; Gluth et al., 2015; Shadlen & Shohamy, 2016). Therefore, our first question was this: Is there evidence that school-age children can already make such decisions with relative ease, as predicted by developmental memory models that suggest an early development of associative memory for details (Shing et al., 2010)?

Second, what kind of developmental differences emerge in value-based decisions and which cognitive components may underlie behavioral differences in a diffusion-modeling analysis? To the extent that (a) access to memory and numeric valuation develops, we may expect increases in the (absolute) drift rate v with increasing age (Ratcliff, Love, Thompson, & Opfer, 2012; Thompson et al., 2016). In other words, quality of evidence (signal-to-noise ratio) during retrieval and integration of values from memory could be higher in adults than in children. To the extent that (b) attitudes to outcomes change from childhood to adulthood, people could sample information about gains and losses differently (v_{bias}) and/or they could have different a-priori response biases to accept or reject presented objects (z/a). Moreover, developmental research has found higher impulsivity and higher tolerance to ambiguity in adolescents than in adults (Li, Brannon, & Huettel, 2015; Steinberg, 2008; Tymula et al., 2012; van den Bos & Hertwig, 2017). Therefore, to the extent that (c) adults require more evidence and sample more information than younger age groups, they could set wider decision boundaries to reduce uncertainty from memory. Finally, younger children may need more time than older children or adults to perceptually encode the presented information and to execute motor responses after a decision is made (e.g., Ratcliff et al., 2012). Therefore, to the extent that (d) basic perceptual and motor abilities develop across childhood, we may expect shorter nondecision latency T_{er} with increasing age.

Our third question addressed the extent to which sources of developmental differences in value-based decisions could be accounted for by other core cognitive abilities. We therefore conducted correlational analyses between the model parameters and individual-difference measures from cognitive and numerical abilities tests.

Method

Participants

We collected data from 30 fourth-grade children (9-10 years), 30 sixth-grade early adolescents (11-12 years) from primary schools, and 30 younger adults (18-30 years; mostly students) during the year 2014. Data from two fourth-graders were lost due to experimenter error. Given these sample sizes and an alpha level of .05, the statistical power to detect medium-sized interaction effects of $\eta^2 = .06$ between age group and experimental manipulations on dependent variables was at least .80. Participant characteristics and their scores from tests of cognitive abilities are reported in Table 1. Participants were paid volunteers of various socioeconomic backgrounds (from predominantly middle and upper class Caucasian families) in the city Berlin, Germany, and were recruited through local advertisements and a database that was only accessible to a limited number of researchers for recruitment purposes at a research institution. In Berlin, pupils finish primary school after the sixth grade. All participants (except one younger adult) were native speakers of German. However, 13% of adults and 23% of children and adolescents also spoke other languages at home. In 7% of adults and 4% of children and adolescents, both parents had a migration background. Participants gave informed consent (for minors this was done by a primary caretaker) and were tested separately in a quiet laboratory room; all procedures were approved by a local ethics committee. Based on pilot testing, all age groups could be expected to understand the task instructions and to deal with the stimulus materials.

Materials and Procedure

Stimuli were four shapes and four colors and the resulting 16 shape-color combinations. All stimuli were presented on a computer screen with a black background. During an initial *training phase*, participants learned to associate four different colors and four different shapes with gains and losses of different magnitude ranges. The training for color and shape symbols proceeded in separate blocks. The mapping of specific symbols to magnitude ranges was randomized (which color or shape represented which magnitude range) and the symbols used for gain and loss domains were counterbalanced across participants (whether colors represented gains and shapes losses, or vice versa). At each trial during the training phase, two symbols appeared on the screen. Participants were asked to choose the color (shape) that represented the larger gain (loss) and received feedback showing specific point values after each response (see Figure 2A). The point values were drawn randomly at each trial from uniform distributions. The underlying distributions (magnitude ranges) for the different symbols within each domain (shapes or colors) implied a rank order: for example, the values for the color green (experienced across trials by a given participant) were never smaller than those for color blue. All participants completed at least three training blocks, each including 12 trials (the 4×3 possible pairings of symbols within a domain). If accuracy was below 90% (i.e., more than one error) after three blocks for the gain and loss domain, respectively, another block was presented until the accuracy criterion was reached.

In the subsequent *decision task*, participants saw on each trial an object representing a combination of a color and a shape (Figure 2B). Participants had to decide whether to accept or reject each object. Specifically, participants were instructed that an object's overall net value could be positive ("gain object") or negative ("loss object") and that they would collect the points associated with accepted objects and thus "win or lose points, based on their decisions—like in computer games in which a score is accumulated". Participants were informed that these points would later be translated into a monetary bonus (up to €5), based on the mappings between colors/shapes and their gain/loss magnitude. Participants were asked to make decisions quickly.

To ensure that all participants understood the instructions, they first worked self-paced through two examples shown on a paper sheet; moreover, a trained experimenter explained the instructions to each participant, who then explained them back in own words. If necessary, this procedure was repeated until the experimenter was certain that all instructions had been understood. Participants then practiced the task for 12 trials with feedback, in which a response had to be given within a time window ranging from 0.2s to 4s after stimulus onset. On each trial, a fixation cross appeared for 0.5s, followed by a colored shape that remained on the screen until the object was accepted or rejected by pressing the *F* and *J* keys (this key-response mapping was counterbalanced across participants). Participants received feedback for both accuracy and speed ("too fast" and "too slow" messages appeared, respectively, for RTs < 0.2s or RTs > 4s).

Participants then completed 24 blocks of 32 trials each (16 combinations resulting from pairing the four colors with the four shapes, shown twice per block, in random sequence). No further trial-wise feedback was provided. During the first block, participants could still ask questions and consult the experimenter; data from the first two blocks were discarded as practice in all subsequent analyses, leaving up to 704 trials available for the analyses of the decision task.

After each block, participants could take breaks and saw a summary with a running score of collected points and their average RT.

At the end of the session, participants received a *memory test*, in which the shapes and colors were again presented. Participants were asked to indicate which of two shapes or colors represented a larger gain or loss (as in the training phase but without getting feedback). Additionally, they completed questionnaires and a cognitive test battery measuring aspects of fluid, numerical, and verbal abilities (further details are in the Supplement).

Results

Training Phase

After the three training blocks, 61% of the fourth-graders, 60% of sixth-graders, and 70% of the adults had reached the accuracy criterion of at least 90% correct comparative judgments (see Table 1 for a detailed summary). A 3 (Age Group) × 2 (Value Domain: gains, losses) ANOVA indicated no significant difference between age groups in the number of blocks required to reach this accuracy criterion, $F(2, 85) = 1.57$, $p = .21$. Stimuli representing gains and losses were learned equally, $F(1, 85) = 1.30$, $p = .26$, and this held irrespective of age group [$F(2, 85) = 1.44$, $p = .24$, for the Age Group × Value Domain interaction]. Accuracy in the final training block, in which participants had reached the 90% criterion, did not differ between age groups, $F(2, 85) = 2.74$, $p = .07$, nor between gain and loss stimuli. The interaction between Age Group and Value Domain was not significant (Fs < 1). Together, these findings suggest that all three age groups had learned the stimulus-value associations similarly well at the end of the training phase.

Decision Task

We report analyses of decision performance in two parts. First, we examine participants' decision accuracy and RTs and then turn to the diffusion-model analyses.

Accuracy and response time. Table 2 reports the results separately for gain and loss objects as a function of the magnitude of their net value (i.e., the difference between the average gain and loss, calculated by using the midpoint of the respective value ranges), represented by four value categories: 0, ±40, ±80, ±120 points (see Appendix A for details). Acceptances of objects with a positive net value (based on the above categorization) were classified as hits and incorrect acceptances of objects with a negative net value as false alarms. The signal-detection discriminability index d', based on the hit rates and false alarm rates, can thus be interpreted as a measure of participants' discrimination ability between objects with positive vs. negative net values.

Overall, discriminability (i.e., d') was well above chance level for all age groups. This suggests that the elementary school children were already capable of performing the rapid value-based task. A 4 × 3 ANOVA with net-value level as within-subjects factor and age group (4th grade, 6th grade, adults) as between-subjects factor indicated that discriminability was lower for the younger age groups, $F(2, 85) = 10.19$, $p < .01$, $\eta^2 = .19$. In addition, discriminability was higher for objects of higher value, $F(3, 255) = 619.37$, Greenhouse-Geisser $\varepsilon = 0.84$, $p < .01$, $\eta^2 = .87$. This effect was more pronounced for the adult participants than for the younger groups, as indicated by a significant interaction between value level and age group, $F(6, 255) = 5.28$, $p < .01$, $\eta^2 = .02$; that is, adults were more sensitive to different net values of the objects than children or adolescents.

For RTs of correct decisions, an ANOVA with value level and domain (gains vs. losses) as within-subjects factors and age group as between-subjects factor indicated that decisions were faster the higher the net value, $F(3, 255) = 66.32$, $\varepsilon = 0.52$, $p < .01$, $\eta^2 = .43$. In addition, decisions were faster for gain than for loss objects, $F(1, 85) = 19.27$, $p < .01$, $\eta^2 = .18$. There was no effect of age group ($F < 1$) and interactions with value domain and value level were not significant (largest $F < 2.28$; ps $> .10$).

Diffusion-model analysis. To disentangle different cognitive components contributing to accuracy and RTs in the decision task, we fit diffusion models (Ratcliff, 1978; Ratcliff & McKoon, 2008) to each individual's data. For parameter estimation, a maximum-likelihood method was used, employing the code from *fast-dm* (Voss et al., 2004). We chose this method because estimation efficiency was an important aspect; nonetheless, the use of other estimation approaches (such as chi-square minimization) led to very similar results (for a discussion of different parameter-estimation approaches, see Ratcliff & McKoon, 2008; Lerche, Voss, & Nagler, 2017).

To strive for a good balance between model fit and model complexity, we first performed a systematic comparison of conceptually plausible diffusion-model variants: In addition to a full model that included all variability parameters η, s_z, s_t (Ratcliff & McKoon, 2008), we also examined several constrained model variants (cf. van Ravenzwaaij, Donkin, & Vandekerckhove, 2017). In each of them, different constellations of parameters were free to vary across the within-subject factors' net value magnitude and value domain. For model comparison, we determined for each variant the BIC that quantifies the trade-off between a model's goodness of fit and the number of parameters (see Appendix B). The goal was to identify the model that performed best in describing decision behavior for the average participant as well as for the majority of participants. The model variants that we considered did not allow parameters a and z to depend on properties of the stimuli. The reason is that these parameters are assumed to be set before stimulus onset and may vary only if a decision maker can anticipate or know which condition is tested (Ratcliff, 1978), which was not the case in the present paradigm.

The best-performing model (which allowed drift rates to vary as a function of domain and net value) was then employed for more detailed investigations of developmental differences in the model parameters. The selected model variant fit the data from all age groups well on the individual level. Further information about qualitative and quantitative tests of the models is in Appendices B and D.

We focus on the model components that are important for comparing the age groups in the present paradigm: the speed of information uptake (drift rate parameter v); the amount of evidence required to make a decision (boundary separation parameter a); nondecision time T_{er}; stimulus-evaluation bias (v_{bias}) and response bias (z/a), that indicate whether participants have specific tendencies of evaluating or responding differently to gains and losses. Comprehensive distribution plots of all parameters are provided in the Supplement.

For all modeling analyses, trials with RTs shorter than 0.25s or longer than 3.5s were discarded (1% of the trials). The percentages eliminated were 0.96%, 1.27%, and 0.72% for fourth-graders, sixth-graders, and college-aged adults, respectively, and did not differ between age groups ($F < 1$). We also examined other cutoff criteria for extreme RTs in the decision task, but they would not have altered our conclusions. A mixed-design 3 × 4 × 2 ANOVA with between-subjects factor age group and within-subject factors net value and value domain indicated that (absolute) drift rate was higher when net value of an object was higher, $F(3, 255) = 459.67$, $\varepsilon = 0.75$, $p < .01$, $\eta^2 = .83$. As shown in Figure 3A and 3B, these increases were approximately linear in all age groups. In line with this, Ratcliff (2014) has shown that drift-rate functions of stimulus difficulty are often linear—even when corresponding psychometric choice functions have a typical sigmoid shape (see also Thompson et al., 2016). Second, the drift rate differed between age groups, $F(2, 85) = 5.50$, $p < .01$, $\eta^2 = .11$, indicating more efficient processing for adults than for the younger age groups. These age differences were more pronounced in the gain than in the loss domain [$F(2, 85) = 4.38$, $p = .02$, $\eta^2 = .09$, for the Age × Domain interaction] and more pronounced for objects with higher net values [as indicated by an interaction between the age group and net value, $F(6, 255) = 4.89$, $\varepsilon = 0.75$, $p < .01$, $\eta^2 = .02$]. Third, the (absolute) drift rates were higher for gain than for loss objects, $F(1, 85) = 9.61$, $p < .01$, $\eta^2 = .09$. These gain–loss differences in drift were more pronounced for objects with higher net values [as indicated by an interaction between value domain and net value, $F(3, 255) = 6.63$, $\varepsilon = 0.66$, $p < .01$, $\eta^2 = .07$]. The three-way interaction with age group was also significant [$F(6, 255) = 2.54$, $\varepsilon = 0.66$, $p = .04$, $\eta^2 = .05$].

Regarding the other diffusion-model parameters (see Figure 4), nondecision time T_{er} was lower in older age groups, $F(2, 85) = 4.47$, $p = .01$, $\eta^2 = .10$, as was variability in nondecision time (s_t), $F(2, 85) = 3.25$, $p = .04$, $\eta^2 = .07$. Notably, there were no effects of age on boundary separation or on the relative position z/a of the starting point (Fs < 1). Moreover, the modeling revealed no response biases (deviations of z/a from .50) in any age group (all ts < 1.27; ps > .21).

Figure 3C plots stimulus-evaluation bias: Overall, there were main effects of age group (indicating that stimulus evaluation was more strongly biased towards gains in adults than in younger age groups), $F(2, 85) = 4.25$, $p = .02$, $\eta^2 = .09$, and of net value (indicating that the bias for gains increased with higher net-value levels), $F(3, 255) = 5.14$, $\varepsilon = .40$, $p < .01$, $\eta^2 = .06$. For objects with gain and loss attributes of equal average magnitude (i.e., with a net-value of zero points) there was no stimulus-evaluation bias, $t(87) < 1$.

In sum, the modeling revealed developmental differences in the speed of information uptake v and in nondecision time T_{er}. Younger adults' more efficient sampling was particularly pronounced for objects of higher positive net value. Age differences in the boundary-separation parameter a and in starting-point position z/a, were negligible. Notably, adults (but not the younger age groups) showed a stimulus-evaluation bias towards gain-related (and not towards loss-related) information. Hence, there was little evidence for loss aversion in any of the age groups. To examine the generality of this finding, we also analyzed traditional choice-based measures of outcome attitude using regression.

Logistic regression. Following Tom et al. (2007), we assessed sensitivity to gains and losses by fitting a logistic regression to each person's decisions (with gain and loss magnitudes, respectively, as predictor variables and object acceptance vs. rejection as criterion). Based on this analysis, we calculated the ratio of loss/gain impact (regression weights) to obtain individual λ parameters (that represent loss aversion in the prospect-theory value function; Tversky & Kahneman, 1992). Analyses of individual λ parameters provided no evidence for loss aversion (median $\lambda < 1$) and otherwise led to similar conclusions as reported above. Moreover, the λ parameters correlated moderately with stimulus-evaluation bias (v_{bias}) in the diffusion model. Further details of these analyses are in the Supplement.

Memory Test

As could be expected, post-task memory for the value-symbol associations at the end of the session (approximately 45 min after the training phase) had declined relative to the accuracy achieved at the end of the training, $F(1, 85) = 33.62$, $p < .01$, $\eta^2 = .28$. There were no age differences in accuracy in the memory test ($F < 1$). However, measurement time (pre vs. post) interacted with value domain, $F\ (1, 85) = 33.62$, $p < .01$, $\eta^2 = .28$, reflecting that although accuracy for gains and losses did not differ at the end of training (before the decision task), after the study all age groups remembered symbol-value mappings somewhat better for gains than for losses, F (1, 85) = 4.86, p = .03, η^2 = .05.

Relation to Cognitive Ability Measures

What are possible reasons for the observed age differences in drift rate and in nondecision time? On the one hand, they could reflect differences in the mnemonic accessibility of feature-value mappings; alternatively, they could reflect differences in how the age groups integrate the relevant information during the decision process to determine net evidence. To explore these possibilities, we analyzed the relation between the model components and measures of core cognitive abilities.

Specifically, we examined whether individual differences in memory, numerical, fluid, and verbal abilities from a cognitive test battery could account for observed age differences in drift rate and nondecision time. Age groups differed across cognitive tests (with the general and expected pattern of higher scores with increasing age; Table 1). Detailed descriptions and zero-order correlations between all test scales and model parameters are in the Supplemental Materials. Moreover, correlations among RTs, accuracy, and model parameters largely replicated the patterns found in many other decision paradigms (Ratcliff & McKoon, 2008). That is, RTs were most strongly related to boundary separation (r = .88) and accuracy to drift rate (r = .60), but there was no relation between drift and boundaries across subjects (r = −.01).

Multiple mediation analysis. We first computed composites of average drift, of memory, numerical, and fluid abilities, by grouping the test scales based on their conceptual similarity (Table 1). For example, the numerical composite included tests of subtraction (one, two, and three-digit numbers), of arithmetic number comparison, and of number-line mapping.

The fluid-abilities composite included tests of cognitive speed (digit-symbol substitution), of logical reasoning (matrices), and forward and backward span. The memory measure included participants' performance during learning and in the post-task memory test.

We next examined which cognitive abilities might account for the observed age differences in drift rate and nondecision time: Notably, multiple mediation analysis (see Appendix C for details) with age as predictor, and memory, fluid, numerical, verbal abilities as mediators, revealed that developmental differences in drift rate were fully mediated through memory and numerical abilities—but not through fluid or verbal ability scores. There were no significant indirect effects of age on nondecision time through any of these measures. These results suggest that developmental differences in value-based decisions are driven by both memory access (decoding the value domain and the magnitude of an object's attributes) and numerical abilities (potentially related to integration and balancing of gains and loss values).

Discussion

How do children make value-based decisions that require retrieval of gain and loss attributes from memory? And which cognitive abilities might explain possible developmental differences? We addressed these questions in a paradigm in which children, young adolescents, and adults, made rapid decisions about objects that differed systematically in value and that had monetary consequences.

The findings show that people in all age groups were faster and more accurate in responding to objects associated with more extreme gain or loss values. Concordant with this, economic models propose that the value difference between attributes determines the difficulty or strength in preferential decisions (e.g., Glimcher & Fehr, 2014; Johnson & Ratcliff, 2014; Konovalov & Krajbich, 2017). The current study confirmed this regularity in children's value-based decision making. Notably, young adolescents as well as children achieved relatively high accuracy and monetary reward, compared to adults. Nonetheless, behavioral and modeling analyses also revealed systematic developmental differences: First, discrimination ability between positive gain and negative loss objects was somewhat lower in children and adolescents than in adults. In the modeling, this was reflected in reduced information uptake during decisions (drift rate v) in the younger age groups and lower sensitivity to different net-value magnitudes, indicating developmental differences in memory access and value integration that emerged particularly for objects of higher positive value. Second, no age group showed a-priori response tendencies to accept or reject objects (response bias z/a), but adults were positively biased during stimulus evaluation (i.e., a relative facilitation of sampling gain-related information). Hence, there was little evidence for loss aversion, but adults differed from the younger age groups in how they extracted information from memory to evaluate objects. Third, there were no developmental differences in decision cautiousness (parameter a), suggesting that age groups required similar levels of evidence to come to a decision and tolerated similar levels of mnemonic uncertainty. Fourth, nondecision latency (parameter T_{er}) decreased with age. This indicates that the speed of elementary perceptual and motor processes develops from childhood to adulthood, replicating a common pattern from developmental modeling studies in other rapid decision tasks (Ratcliff et al., 2012; Thompson et al., 2016). Finally, the observed age differences in information uptake (drift v) were fully mediated through measures of participants' numerical abilities and their memory of symbol-value mappings, while controlling for fluid and verbal abilities.

Implications for Understanding the Development of Value-Based Decisions

Based on research with adults (Gluth et al., 2015; Shadlen & Shohamy, 2016), we assumed that value-based decisions are closely intertwined with associative remembering. Lifespan theories of memory development propose that associative memory components are relatively mature by middle childhood and involve early developed brain structures (e.g., Shing et al., 2010). The present findings confirmed this expectation for value-based choice in children and young adolescents, who showed similar performance: Decisions were remarkably accurate and rapid, in line with the assumption that the mnemonic abilities required for value-based choice (such as forming associations between object features and their values) are developed from an early age. These results also fit the notion that children, within their cognitive capabilities, tend to rely on memory for perceptual detail (Brainerd, Holliday, & Reyna, 2004). For example, to the extent that decisions involve specific quantitative cues about value, fuzzy-trace theory suggests that younger children may already utilize relatively precise verbatim memory representations (Reyna, 2012; Reyna et al., 2003).

In the current study, adults assigned relatively more weight to gain than to loss attributes during stimulus evaluation whereas neither children nor young adolescents showed such positivity effects. Higher motivational sensitivity to gains in adults than children has also been observed in other reward-based paradigms and has been attributed to the development of dopaminergic neuromodulation (see Hämmerer et al., 2010). However, from the perspective of neurodevelopmental imbalance models, a more pronounced impact of gains in adolescents than in adults was expected (e.g., Somerville et al., 2010). The adolescents in the present study were relatively young (*M* = 11.6 years). It is conceivable that increased responsiveness to rewarding stimuli emerges only during later adolescence, and particularly when decisions involve active exploration in heightened affective settings (Figner et al., 2009; Rosenbaum et al., 2017).

Notably, there was little indication of loss aversion in any age group. Our modeling analysis of stimulus evaluation (v_{bias}) and response bias (z/a) did not indicate stronger impact of loss than gain attributes on decisions. In other words, to accept objects in the task, people did not require that remembered gain values were larger than loss values; moreover, people did not require more evidence for acceptance than for rejection of objects. This replicates and extends recent research, suggesting that loss aversion may not operate in children, adolescents, and adults when gain and loss outcomes appear concurrently, or when losses are experienced repeatedly rather than described (Erev et al., 2008; Yechiam & Hochman, 2013). Other research also suggests that adults may even seek for gains when smaller amounts of money are at stake (e.g., Clay et al., 2017; Harinck et al., 2007).

Age differences in the quality of value-based decisions (drift rate v) emerged mainly for objects of high positive (but not of negative) value in the current study. This age effect on decision quality was fully mediated by people's memory for symbol-value mappings (before and after the decision task) and their numerical abilities. This could reflect developmental differences in the mnemonic accessibility of associated gain values or in the numeric integration of gains and losses. For example, adults might have focused more than younger age groups on gain symbols during encoding (training), leading to facilitated access to this information during the subsequent decisions. Although it is for future research to disentangle these possibilities in greater detail, we could exclude fluid abilities and cognitive speed as mediators of age effects on decision performance. Fluid cognitive abilities could be more relevant in decisions with more strategic requirements that are often made at slower pace, for example, in decisions among attributes in information boards (Mata, von Helversen, & Rieskamp, 2011) or among risky lotteries, in which probability and value information is explicitly described (Defoe et al., 2015).

There are some limitations to the generalizability of our findings. For instance, we used a selection of perceptual stimuli whose arbitrary associations with values had to be learned in an initial study phase. Whereas this approach provides good experimental control over prior knowledge and stimulus features, it is important to extend the research on cost-benefit decisions to other environments, in which children might use natural associations between real-world objects (e.g., toys, snacks, or other familiar items) and subjective value, based on individual learning histories (e.g., Horn, Ruggeri, & Pachur, 2016). It is also an open question whether the current findings generalize to other non-monetary forms of cost and benefits (beyond point values, tokens, and monetary payoffs). In future research, it would also be interesting to investigate whether our findings about core cognitive constructs extend to other value-based decisions that are not memory based or require different memory representations. For example, one approach could be to systematically vary the memory demands (e.g., decisions from description vs. decisions from memory) and magnitude information (e.g., Arabic numerals vs. icons of different complexity) to assess the role of memory and numerical abilities as a function of task characteristics. Moreover, the systematic investigation of rapid value-based decisions that tap different memory representations (e.g., gist or verbatim representations) could be an interesting avenue: Developmental differences in performance may depend on whether decision tasks elicit meaning-based, qualitative or quantitative interpretation of the outcomes (e.g., Reyna, 2012).

In conclusion, the current study provides initial insight into the development of value-based decisions from memory and shows that developmental differences are associated with separable cognitive abilities. Our findings reveal that school-age children can already make surprisingly good and rapid decisions about value and highlight the importance of basic associative memory and numeric processes to explain developmental differences in value-based decisions.

References

Basten, U., Biele, G., Heekeren, H. R., & Fiebach, C. J. (2010). How the brain integrates costs and benefits during decision making. *Proceedings of the National Academy of Sciences, 107,* 21767–21772.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition, 50,* 7–15.

Beitz, K. M., Salthouse, T. A., & Davis, H. P. (2014). Performance on the Iowa Gambling Task: From 5 to 89 years of age. *Journal of Experimental Psychology: General, 143,* 1677–1689.

Boyer, T. W. (2006). The development of risk-taking: A multi-perspective review. *Developmental Review, 26,* 291–345.

Brainerd, C. J., Holliday, R. E., & Reyna, V. F. (2004). Behavioral measurement of remembering phenomenologies: So simple a child can do it. *Child Development, 75,* 505–522.

Busemeyer, J. R., & Diederich, A. (2002). Survey of decision field theory. *Mathematical Social Sciences, 43,* 345–370.

Cauffman, E., Shulman, E. P., Steinberg, L., Claus, E., Banich, M. T., Graham, S., & Woolard, J. (2010). Age differences in affective decision making as indexed by performance on the Iowa Gambling Task. *Developmental Psychology, 46,* 193–207.

Clay, S. N., Clithero, J. A., Harris, A. M., & Reed, C. L. (2017). Loss aversion reflects information accumulation, not bias: A drift-diffusion model study. *Frontiers in Psychology, 8,* 1708.

Costantini, A. F., & Hoving, K. L. (1973). The effectiveness of reward and punishment contingencies on response inhibition. *Journal of Experimental Child Psychology, 16,* 484– 494.

Defoe, I. N., Dubas, J. S., Figner, B., & van Aken, M. A. (2015). A meta-analysis on age differences in risky decision making: Adolescents versus children and adults. *Psychological Bulletin, 141,* 48–84.

Enkavi, A. Z., Weber, B., Zweyer, I., Wagner, J., Elger, C. E., Weber, E. U., & Johnson, E. J. (2017). Evidence for hippocampal dependence of value-based decisions. *Scientific Reports,* 7, 17738.

Erev, I., Ert, E., & Yechiam, E. (2008). Loss aversion, diminishing sensitivity, and the effect of experience on repeated decisions. *Journal of Behavioral Decision Making*, *21*, 575–597.

Figner, B., Mackinlay, R. J., Wilkening, F., & Weber, E. U. (2009). Affective and deliberative processes in risky choice: Age differences in risk taking in the Columbia Card Task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 709–730.

Fry, A. F., & Hale, S. (2000). Relationships among processing speed, working memory, and fluid intelligence in children. *Biological Psychology, 54,* 1–34.

Gächter, S., Johnson, E.J., & & Herrmann, A. (2007). Individual-level loss aversion in riskless and risky choices. *IZA Discussion Paper* No. 2961. Technical report.

Glimcher, P. W., & Fehr, E. (2014). *Neuroeconomics.* New York: Academic Press.

Gluth, S., Sommer, T., Rieskamp, J., & Büchel, C. (2015). Effective connectivity between hippocampus and ventromedial prefrontal cortex controls preferential choices from memory. *Neuron, 86,* 1078–1090.

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences, 109,* 11116–11120.

Hämmerer, D., Li, S.-C., Müller, V., & Lindenberger, U. (2010). Lifespan differences in electrophysiological correlates of monitoring gains and losses during probabilistic reinforcement learning. *Journal of Cognitive Neuroscience, 23,* 579–592.

Harbaugh, W. T., Krause, K., & Vesterlund, L. (2001). Are adults better behaved than children? Age, experience, and the endowment effect. *Economics Letters, 70,* 175–181.

Harinck, F., Van Dijk, E., Van Beest, I., & Mersmann, P. (2007). When gains loom larger than losses: Reversed loss aversion for small amounts of money. *Psychological Science, 18,* 1099–1105.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* New York: Guilford Press.

Horn, S. S., Bayen, U. J., & Smith, R. E. (2013). Adult age differences in interference from a prospective-memory task: a diffusion model analysis. *Psychonomic Bulletin & Review, 20,* 1266–1273.

Horn, S. S., Ruggeri, A., & Pachur, T. (2016). The development of adaptive decision making: Recognition-based inference in children and adolescents. *Developmental Psychology, 52,* 1470–1485.

Johnson, E. J., & Ratcliff, R. (2014). Computational and process models of decision making in psychology and behavioral economics. In P. W. Glimcher & E. Fehr (Eds.), *Neuroeconomics* (pp. 35–47). New York: Academic Press.

Kail, R. (1991). Developmental change in speed of processing during childhood and adolescence. *Psychological Bulletin, 109,* 490–501.

Konovalov, A., & Krajbich, I. (2017). Revealed indifference: Using response times to infer preferences. *Social Science Research Network.* doi.org/10.2139/ssrn.3024233

Lerche, V., Voss, A., & Nagler, M. (2017). How many trials are required for parameter estimation in diffusion modeling? A comparison of different optimization criteria. *Behavior Research Methods, 49,* 513–537.

Lewin, K. (1935). *A dynamic theory of personality.* New York: McGraw-Hill.

Li, R., Brannon, E. M., & Huettel, S. A. (2015). Children do not exhibit ambiguity aversion despite intact familiarity bias. *Frontiers in Psychology, 5,* 1519.

Mata, R., von Helversen, B., & Rieskamp, J. (2011). When easy comes hard: The development of adaptive strategy selection. *Child Development, 82,* 687–700.

Rangel, A., Camerer, C., & Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nature Reviews Neuroscience, 9,* 545–556.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review, 85,* 59–108.

Ratcliff, R. (2014). Measuring psychometric functions with the diffusion model. *Journal of Experimental Psychology: Human Perception and Performance, 40,* 870–888.

Ratcliff, R., Love, J., Thompson, C. A., & Opfer, J. E. (2012). Children are not like older adults: A diffusion model analysis of developmental changes in speeded responses. *Child Development, 83,* 367–381.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation, 20,* 873–922.

Reyna, V. F. (2012). A new intuitionism: Meaning, memory, and development in fuzzy-trace theory. *Judgment and Decision Making, 7,* 332–359.

Reyna, V. F., & Brainerd, C. J. (1994). The origins of probability judgment: A review of data and theories. In G. Wright & P. Ayton (Eds.), *Subjective probability* (pp. 239–272). Chichester, United Kingdom: Wiley.

Reyna, V. F., Lloyd, F. J., & Brainerd, C. J. (2003). Memory, development, and rationality: An integrative theory of judgment and decision making. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on judgment and decision research* (pp. 203-245). Cambridge: University Press.

Rosenbaum, G. M., Venkatraman, V., Steinberg, L., & Chein, J. M. (2017). The influences of described and experienced information on adolescent risky decision making. *Developmental Review.*

Schley, D. R., & Peters, E. (2014). Assessing "economic value": Symbolic-number mappings predict risky and riskless valuations. *Psychological Science, 25,* 753–761.

Schlottmann, A., & Wilkening, F. (2011). Judgment and decision making in young children. In M. K. Dhami, A. Schlottmann, & M. R. Waldmann (Eds.), *Judgment and decision making as a skill: Learning, development, and evolution* (pp. 55–83). Cambridge, United Kingdom: Cambridge University Press

Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron, 90,* 927–939.

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Müller, V., Li, S. C., & Lindenberger, U. (2010). Episodic memory across the lifespan: The contributions of associative and strategic components. *Neuroscience and Biobehavioral Reviews, 34,* 1080–1091.

Siegler, R. S. (2016). Magnitude knowledge: the common core of numerical development. *Developmental Science, 19,* 341–361.

Slovic, P., & Lichtenstein, S. (1968). Relative importance of probabilities and payoffs in risk taking. *Journal of Experimental Psychology, 78,* 1–18.

Sluzenski, J., Newcombe, N. S., & Kovacs, S. L. (2006). Binding, relational memory, and recall of naturalistic events: A developmental perspective. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 32,* 89–100.

Somerville, L. H., Jones, R. M., & Casey, B. J. (2010). A time of change: behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. *Brain and Cognition, 72,* 124–133.

Spaniol, J., Voss, A., & Grady, C. L. (2008). Aging and emotional memory: cognitive mechanisms underlying the positivity effect. *Psychology and Aging, 23,* 859–872.

Steinberg, L. (2008). A social neuroscience perspective on adolescent risk-taking. *Developmental Review, 28,* 78–106.

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive Psychology, 53,* 1–26.

Thompson, C. A., Ratcliff, R., & McKoon, G. (2016). Individual differences in the components of children's and adults' information processing for simple symbolic and non-symbolic numeric decisions. *Journal of Experimental Child Psychology, 150,* 48–71.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science, 315,* 515–518.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics, 106,* 1039–1061.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5,* 297–323.

Tymula, A., Belmaker, L. A. R., Roy, A. K., Ruderman, L., Manson, K., Glimcher, P. W., & Levy, I. (2012). Adolescents’ risk-taking behavior is driven by tolerance to ambiguity. *Proceedings of the National Academy of Sciences, 109,* 17135–17140.

van den Bos, W., & Hertwig, R. (2017). Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making. *Scientific Reports, 7,* 40962.

van Duijvenvoorde, A .C. K., Jansen, B. R. J. & Huizenga, H. M. (2015). Risky choice from childhood to adulthood: changes in decision strategies, affect, and control. In E.A. Wilhelms & V.F. Reyna (Eds.), *Neuroeconomics, judgment, and decision making* (Frontiers of cognitive psychology) (pp. 203-218). New York: Psychology Press.

van Duijvenvoorde, A. C., Huizenga, H. M., Somerville, L. H., Delgado, M. R., Powers, A., Weeda, W. D., Casey, B. J., Weber, E. U. & Figner, B. (2015). Neural correlates of expected risks and returns in risky choice across development. *The Journal of Neuroscience, 35,* 1549–1560.

van Duijvenvoorde, A. C., Zanolie, K., Rombouts, S. A., Raijmakers, M. E., & Crone, E. A. (2008). Evaluating the negative or valuing the positive? Neural mechanisms supporting feedback-based learning across development. *The Journal of Neuroscience, 28,* 9495–9503.

van Ravenzwaaij, D., Donkin, C., & Vandekerckhove, J. (2017). The EZ Diffusion Model Provides a Powerful Test of Simple Empirical Effects. *Psychonomic Bulletin & Review, 24,* 547–556.

Voss, A., Rothermund, K., & Voss, J. (2004). Interpreting the parameters of the diffusion model: An empirical validation. *Memory & Cognition, 32,* 1206–1220.

Weber, E. U., & Johnson, E. J. (2006). Constructing preferences from memories. In S. Lichtenstein & P. Slovic (Eds.), *The construction of value* (pp. 397–410). New York: Cambridge University Press.

Weller, J. A., Levin, I. P., & Denburg, N. L. (2011). Trajectory of risky decision making for potential gains and losses from ages 5 to 85. *Journal of Behavioral Decision Making, 24,* 331–344.

Yechiam, E., & Hochman, G. (2013). Losses as modulators of attention: Review and analysis of the unique effects of losses over gains. *Psychological Bulletin, 139,* 497–518.

Table 1. *Participant Characteristics, Training Performance, and Cognitive Test Scores*

Variable	Age Group			Age Group			Effect Size [a]
	4th grade	6th grade	Adults	4th grade	6th grade	Adults	
	*M*s			*SD*s			
Age (years)	9.21	11.60	23.50	.42	.50	2.74	
N	28	30	30	–	–	–	
Gender (*N* male)	20	14	17	–	–	–	
Training Performance							
Acc. overall gains	.86	.87	.91	.10	.12	.07	.05
Acc. overall losses	.84	.87	.89	.11	.09	.08	.05
Acc. final gains	.96	.97	.97	.04	.04	.04	.04
Acc. final losses	.96	.97	.97	.04	.04	.04	.03
n blocks gains	3.50	3.87	3.03	1.07	2.16	.18	.06
n blocks losses	4.00	3.60	3.53	2.21	1.45	1.11	.02
Post-Task Memory Test							
Acc. Gains	.92	.91	.95	.08	.09	.09	.03
Acc. Losses	.88	.87	.90	.12	.19	.21	.01
Cognitive Tests [b]							
Vocabulary [c]	21.59	25.00	28.40	2.71	2.89	1.13	.58***
Speed Test (Digit-Symbol) [d]	46.56	60.20	89.83	10.75	8.52	13.66	.73***
Reasoning Matrices [d]	9.59	12.40	12.10	2.71	1.57	1.79	.27***
Digit Span (forward) [d]	8.89	9.17	10.43	1.97	2.28	2.08	.10*
Digit Span (backward) [d]	5.85	6.33	8.07	1.73	1.77	2.08	.21***
Number Line Task [e]	10.19	10.93	11.53	1.62	1.46	.94	.14**
Subtractions [e]	27.30	32.27	36.23	4.54	3.77	1.92	.52***
Arithmetic Comparisons [e]	28.41	34.63	39.00	6.43	4.16	1.14	.50***
Panamath (RT) [e]	0.96	0.86	0.64	0.17	0.14	0.15	.44***

Note. [a] R^2 for main effect of age group; [b] raw test scores; further details about the cognitive tests are in the online supplement; *N* = sample size; grouping of test scales: [c] verbal, [d] fluid, [e] numerical abilities
Acc. = proportion of accurate choices (comparative judgments); overall = across all training blocks; final = in final training block; *n* blocks = number of training blocks needed to reach the learning criterion; RT = response time in *s*

* $p < .05$; ** $p < .01$; *** $p < .001$

Table 2. *Performance in the Decision Task as a Function of Age Group and Value Level: Hit Rate, False-Alarm Rate, d', and Response Times*

Variable	Means			Standard Deviations		
	4th grade	6th grade	Adults	4th grade	6th grade	Adults
N_{trials}	697	695	699	13	17	12
Hit Rate						
HR_1	.43	.51	.56	.24	.21	.22
HR_2	.73	.77	.85	.17	.21	.15
HR_3	.89	.88	.96	.12	.15	.06
HR_4	.96	.93	.98	.07	.17	.02
False Alarm Rate						
FAR_1	.46	.51	.54	.23	.21	.23
FAR_2	.26	.22	.18	.18	.15	.17
FAR_3	.14	.11	.05	.17	.14	.09
FAR_4	.06	.05	.03	.07	.07	.06
Discriminability d'						
d'_1	−.05	−.02	.05	.26	.25	.21
d'_2	1.62	1.89	2.39	.67	1.08	1.01
d'_3	2.81	3.11	4.17	.95	1.34	.98
d'_4	3.71	3.71	4.24	.78	.99	.55
RT gains						
RT_1	791	833	795	238	262	295
RT_2	779	790	734	183	232	195
RT_3	745	727	662	197	184	180
RT_4	697	659	587	128	129	135
RT losses						
RT_1	851	819	855	263	251	309
RT_2	834	810	812	268	250	285
RT_3	792	769	748	218	226	254
RT_4	743	695	686	206	169	226

Note. Indices 1 to 4 refer to the magnitude of the average net value, used for grouping the objects presented in the decision task (see Appendix A): level 1 = 0 points; level 2 = ±40 points; level 3 = ±80 points; level 4 = ±120 points. N_{trials} = Number of trials available for analyses; RT = response time of correct decisions in ms (based on the individual median RTs). Because for objects with large net values some adults committed few errors, we refrained from separate RT analyses of error decisions. However, both correct and error RTs were considered for the modeling. Model parameters are estimated from those conditions where accuracy is not at ceiling (or floor) and then constrain estimates for those conditions where accuracy is high (or low). That is, the model may capture differences in drift rates because other available conditions allow these parameters to be estimated; because RTs differ across the high-accuracy conditions, RTs are sufficient to determine drift rates (see Ratcliff, 2014).

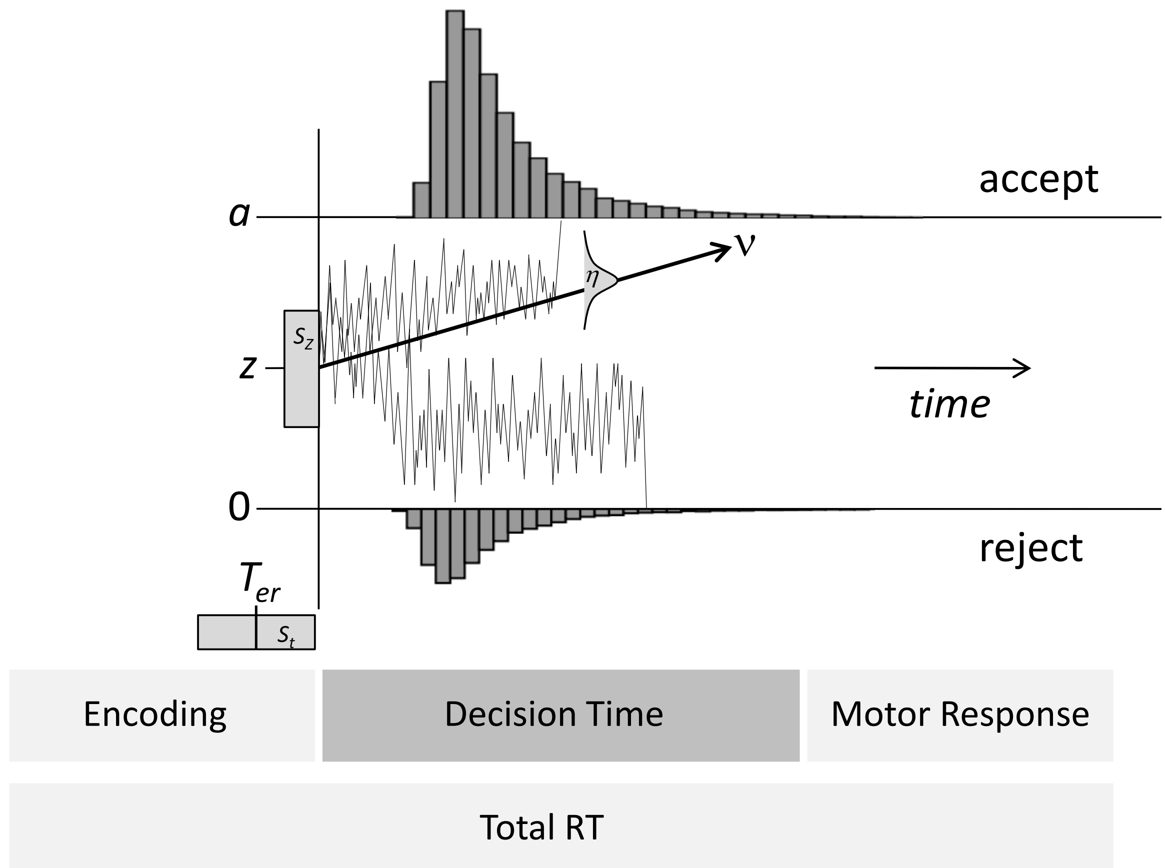


Figure 1. Illustration of the drift-diffusion model. The vertical axis is the decision-related strength-of-evidence axis, and the horizontal axis is the time axis. Diffusion processes start at point *z* and move over time until the upper boundary (positioned at *a*) or the lower boundary (positioned at zero) is reached. Decision time distributions associated with the upper and the lower boundaries are shown. The two oscillating sample tracks illustrate that different boundaries can be reached with the same (in the example: positive) drift rate due to random influence, as given by the diffusion constant *s*. Total RT is the sum of the decision time and a nondecision component T_{er} that measures duration of peripheral processes such as encoding and motor response execution. In all present analyses, the upper boundary was associated with “accept” decisions and the lower boundary with “reject” decisions; *s* was fixed at 1. Adapted with permission from “Adult age differences in interference from a prospective-memory task: a diffusion model analysis” by S. Horn, U. J. Bayen, and R. E. Smith, 2013, Psychonomic Bulletin & Review, p. 1267. © Psychonomic Society.

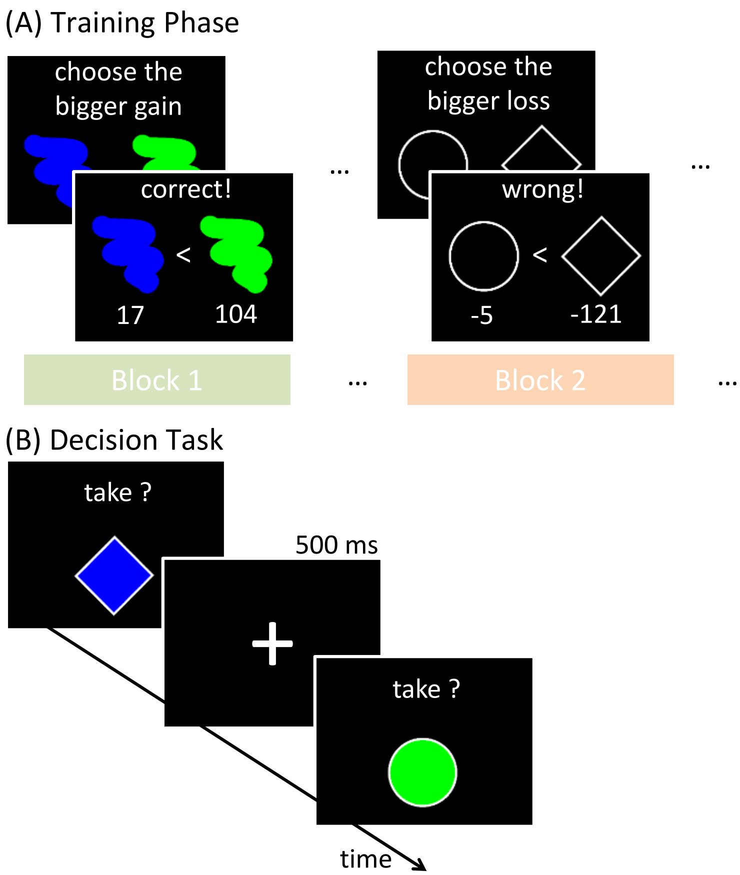


Figure 2. Training procedure and the decision task.

(A) The training phase comprised at least three learning blocks for colors and for shapes, respectively, in alternating sequence. On each trial, participants saw pairs of symbols (e.g., blue and green) and were asked to choose the item with the higher value. After choice, participants received feedback, thereby implicitly learning value ranges associated with specific colors and shapes. For example, the color green might be associated with a gain between +80 and +120 points, drawn from a distribution $U\,\{80, 120\}$. Training was terminated after a criterion (90% accurate comparative judgments) was reached for each domain (colors, shapes). Each learning block included all possible pairings of symbols and each symbol in a pair was shown once on the left and right side of the screen, respectively.

(B) Decision task, in which participants saw feature combinations that were associated with net values. For example, the expected net value of a green circle would result from the combination of the gain value associated with color green (e.g., between + 80 and + 120) and the loss value associated with the circle (e.g., between 0 and −40). Successful gain–loss integration from memory would lead participants to accept this object. An exemplary payoff matrix for a given participant as a function of feature combinations is in Appendix A.

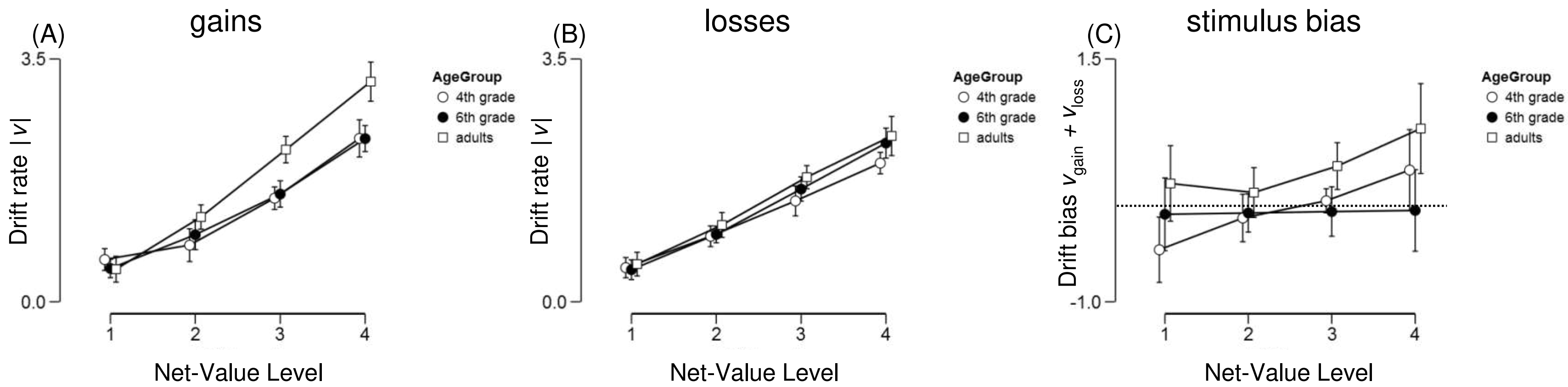


Figure 3. Mean drift-rate parameters of the diffusion model as a function of age group and magnitude of the net value (level 1 = 0 points; level 2 = ±40 points; level 3 = ±80 points; level 4 = ±120 points). Error bars represent 95% confidence intervals.

(A) Absolute drift rates for positive gain objects

(B) Absolute drift rates for negative loss objects

(C) Stimulus-evaluation bias ($v_{gain} + v_{loss}$). The dashed zero line indicates unbiased evaluation (i.e., the same drift for gains and losses)

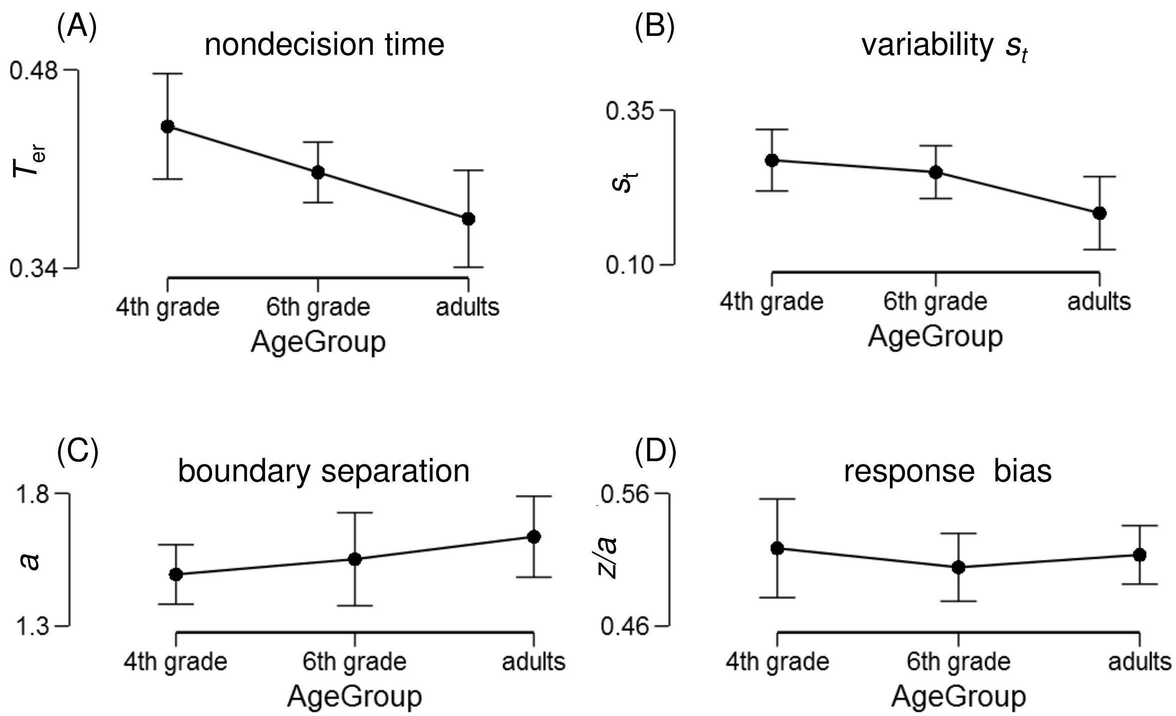


Figure 4. Mean diffusion model parameters as a function of age group (error bars represent 95% confidence intervals).

(A) Nondecision time T_{er} (in s)

(B) Across-trial variability in nondecision time s_t

(C) Boundary separation a (decision cautiousness)

(D) Response bias (relative starting-point position z/a)

Appendix A

Pay-Offs in the Decision Task

Figure A1 shows an exemplary payoff matrix as a function of symbol combinations for a given participant. Color and shape symbols were selected from a larger pool and the mappings of specific symbols to specific value ranges were randomized across participants. This randomization ensured that identical stimulus features differed in values across participants (e.g., the color green could be associated with a high value for one participant, but a low value for another) and that the values of stimuli were not confounded with their specific perceptual or physical properties.

On each trial (during both training and the decision task), an integer value for a feature was drawn from a discrete uniform distribution $U\ \{a, b\}$ with absolute range of $|a-b| = 40$ points and with lower limit $|a| \in (0, 40, 80, 120)$ and upper limit $|b| \in (40, 80, 120, 160)$. During training, each participant learned to associate four colors with four different value ranges and four shapes with four different ranges of monetary costs (or vice versa, counterbalanced across subjects). In the subsequent value-based decision task, each object was a *combination* of both color and shape, and thus carried information about gains and losses, as the two examples (i.e., green circle and blue rectangle) illustrate. Altogether, 16 different objects appeared during the decision task and each was shown twice in each of the experimental blocks. The average (expected) magnitudes of the net value of these objects were −120, −80, −40, 0, +40, +80, +120 points. Objects with a net value of zero (i.e., objects for which, on average, gains and losses, cancel each other out) were included in the task as ambiguous guessing stimuli, for which the diffusion model predicts chance level performance (the feature combinations on the diagonal of the matrix). As expected, accuracy or drift rates did neither differ between age groups nor from chance level for these objects. Nonetheless, these stimuli were included in the modeling and can serve as reference in revealing participants' decision tendencies and for constraining other parameter estimates.

For analyzing participants' performance in the value-based decision task, we classified objects based on their average net value into four magnitude categories (0, ±40, ±80, ±120 points) for gains and losses, respectively. These categories had similar effects on performance as stimulus-difficulty levels in other decision tasks (e.g., the information extractable from a stimulus in a mnemonic or perceptual decision task). For ease of reference, objects with a net value of zero that had point values > 0 on specific trials were classified as "gain objects" and objects with point values < 0 were classified as "loss objects"; for objects with an exact value of zero on specific trials, the domain assignment for analyses was determined randomly with equal probability.

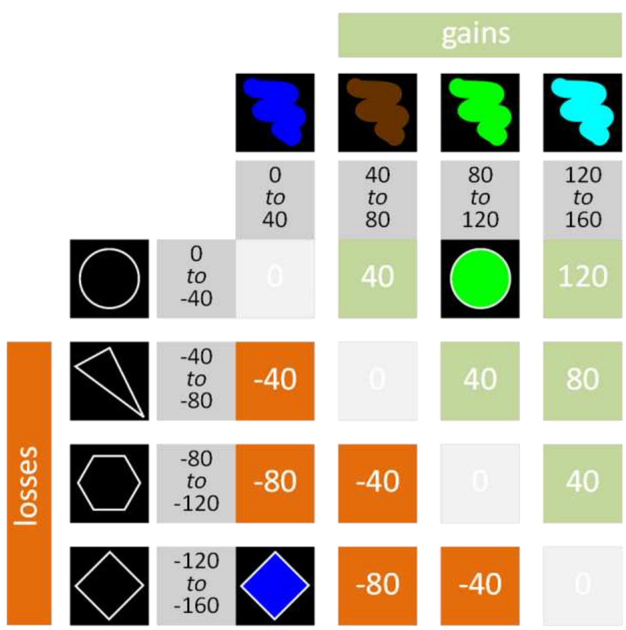


Figure A1. Exemplary payoff matrix for the acceptance of stimuli in the decision task as a function of color-shape feature combinations (for rejection of stimuli, the associated outcome value was generally zero).

Appendix B

Model Comparison

Comparison of diffusion models with different constellations of free parameters. The best-performing model (#19) that was used for further analyses in this study is marked in boldface. The column entries for model parameters v, T_{er}, $B = z/a$, η, s_z, s_t, indicate whether these parameters were constrained (–) or free to vary across the experimental factors (NV = net-value level; D = domain gain/loss), or fixed at specific values (.50 or 0); K = number of estimated model parameters; model-selection indices refer to the likelihood of a model ($-2 \ln L$), the average BIC ranking across participants (separated by age group), and number of participants (n) for which a specific model had the lowest BIC; for model variants 25-28 that assumed equivalent drift v for gain and loss stimuli, data were collapsed across gain and loss stimuli (with the upper and lower boundaries associated with correct and error decisions, respectively). In the selected model, the across-trials variabilities $\eta = s_z = 0$ were fixed (in line with the finding that correct and error RTs were similar) but variability in nondecision time s_t was included (in line with the notion that s_t is often important for capturing the shape of the RT distributions; e.g., Lerche et al., 2017). Finally, we note that analyses with the runner-up BIC model (#23; which additionally assumed unbiased responding; $B = .50$) would have otherwise led to identical conclusions as discussed in the results section.

Table B1 *Results of the Model Comparison*

Model Variant	Model Parameters							K	Model Selection Indices				
#	v	T_{er}	a	B	η	s_z	s_t		$-2 \ln L$	BIC rank			BIC winner n
										4th grade	6th grade	adults	
1	NV,D	NV,D	–	–	–	–	–	21	581.30	11.93	14.20	17.53	0
2	NV,D	NV,D	–	–	–	0	0	19	641.72	21.05	21.07	21.97	0
3	NV,D	NV,D	–	–	0	0	–	19	591.77	10.95	13.23	16.20	0
4	NV,D	NV,D	–	–	0	0	0	18	646.54	18.82	20.73	21.13	0
5	NV,D	NV,D	–	.50	–	–	–	20	594.31	14.89	14.37	17.10	0
6	NV,D	NV,D	–	.50	–	0	0	18	650.07	22.57	20.97	21.03	0
7	NV,D	NV,D	–	.50	0	0	–	18	606.10	13.25	13.80	16.37	0
8	NV,D	NV,D	–	.50	0	0	0	17	656.67	20.86	20.67	20.90	0
9	NV,D	NV	–	–	–	–	–	17	588.84	6.96	8.67	10.63	0
10	NV,D	NV	–	–	–	0	0	15	660.35	18.68	18.83	15.80	0
11	NV,D	NV	–	–	0	0	–	15	599.27	5.75	7.77	9.47	1
12	NV,D	NV	–	–	0	0	0	14	665.67	16.36	18.53	15.07	0
13	NV,D	NV	–	.50	–	–	–	16	606.88	10.32	11.47	10.80	1
14	NV,D	NV	–	.50	–	0	0	14	671.33	20.54	20.63	16.20	0
15	NV,D	NV	–	.50	0	0	–	14	617.07	9.04	9.53	10.03	3
16	NV,D	NV	–	.50	0	0	0	13	676.45	18.64	19.47	16.00	0
17	NV,D	–	–	–	–	–	–	14	597.26	4.89	4.97	5.27	9
18	NV,D	–	–	–	–	0	0	12	683.63	17.50	16.73	13.83	3
19*	**NV,D**	**–**	**–**	**–**	**0**	**0**	**–**	**12**	**605.61**	**2.93**	**3.50**	**3.83**	**22**
20	NV,D	–	–	–	0	0	0	11	687.68	15.39	16.63	12.63	2
21	NV,D	–	–	.50	–	–	–	13	613.08	6.61	5.70	6.03	4
22	NV,D	–	–	.50	–	0	0	11	690.63	18.79	17.37	13.87	0
23	NV,D	–	–	.50	0	0	–	11	623.59	5.32	4.87	4.97	21
24	NV,D	–	–	.50	0	0	0	10	698.98	17.18	17.40	12.97	3
25	NV	–	–	.50	–	–	–	9	705.49	15.00	11.93	16.73	4
26	NV	–	–	.50	–	0	0	7	792.93	24.36	21.20	21.57	0
27	NV	–	–	.50	0	0	–	7	717.28	14.00	11.20	16.57	15
28	NV	–	–	.50	0	0	0	6	796.63	23.43	20.57	21.50	0

Note. *Selected diffusion model variant

Appendix C

Correlational Analyses

Path analytic methods were used to concurrently explore the mediating role of different cognitive abilities in the development of value-based decisions. In the present paradigm, developmental differences emerged mainly in decision accuracy. In the diffusion model analysis, this was reflected in corresponding differences in drift rate v (for positive gain objects) and, to a smaller extent, in nondecision time T_{er}. We used multiple mediation analysis (with cognitive-ability composites as *parallel mediators*; see Figure C1) to predict participants' decision performance (criterion variable Y). The comprehensive results of these analyses (including standardized regression weights with their confidence intervals) are in Table C1.

Age effects on drift rate v (gains) as criterion were fully mediated through numerical ability scores and memory for value-feature mappings (composite of encoding accuracy and post-task memory), but not through fluid ability or verbal test scores. The same pattern emerged for decision accuracy as criterion variable. In contrast, there was no mediation of age through any of these cognitive measures on nondecision time T_{er} (or RTs) as criterion. Finally, we note that mediational path analyses with cross-sectional data might differ from those with longitudinal data (e.g., Lindenberger, von Oertzen, Ghisletta, & Hertzog, 2011).

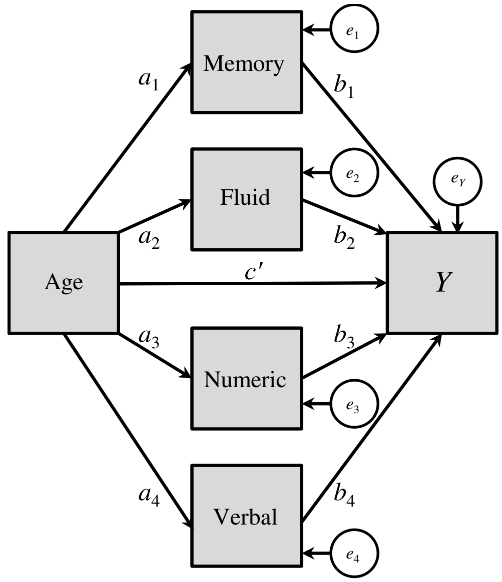


Figure C1. Multiple mediator model that predicts effects of age on decision performance (*Criterion Variable Y*) through measures of mnemonic accuracy (*Memory*), fluid abilities (*Fluid*), numerical/arithmetic abilities (*Numeric*), and vocabulary (*Verbal*).

Table C1

Mediation Analyses with Drift, Nondecision Time, Accuracy, Response Time as Criterion

	Criterion				
	Drift v (gains)	Drift v (losses)	Nondecision Time T_{er}	Accuracy	RT
Model Coefficient/ Description					
Total and direct effects of Age					
c (total effect Age → Y)	**.38 [.18, .58]**	.15 [−.06, .36]	**−.27 [−.48, −.06]**	**.37 [.17, .57]**	−.04 [−.25, .18]
c' (direct effect Age → Y)	.04 [−.22, .30]	−.16 [−.45, .13]	−.25 [−.58, .07]	.09 [−.14, .33]	−.01 [−.35, .32]
Indirect effects (Mediation)					
$a_1 \times b_1$ (through memory)	**.08 [.01, .21]**	**.08 [.01, .21]**	.01 [−.03, .09]	**.11 [.01, .27]**	.02 [−.02, .12]
$a_2 \times b_2$ (through fluid)	.01 [−.19, .18]	.03 [−.17, .21]	.04 [−.15, .23]	.03 [−.13, .17]	.02 [−.18, .24]
$a_3 \times b_3$ (through numeric)	**.20 [.09, .33]**	**.16 [.01, .31]**	−.03 [−.20, .12]	**.17 [.05, .30]**	−.03 [−.20, .16]
$a_4 \times b_4$ (through verbal)	.05 [−.17, .22]	.05 [−.14, .24]	−.03 [−.19, .16]	−.03 [−.20, .11]	−.04 [−.24, .16]
Age → Mediator					
a_1 (age → memory)	**.21 [.01, .42]**	**.21 [.01, .42]**	**.21 [.01, .42]**	**.21 [.01, .42]**	**.21 [.01, .42]**
a_2 (age → fluid)	**.66 [.50, .82]**	**.66 [.50, .82]**	**.66 [.50, .82]**	**.66 [.50, .82]**	**.66 [.50, .82]**
a_3 (age → numeric)	**.60 [.43, .77]**	**.60 [.43, .77]**	**.60 [.43, .77]**	**.60 [.43, .77]**	**.60 [.43, .77]**
a_4 (age → verbal)	**.68 [.52, .84]**	**.68 [.52, .84]**	**.68 [.52, .84]**	**.68 [.52, .84]**	**.68 [.52, .84]**
Mediator → Criterion					
b_1 (memory → Y)	**.38 [.19, .56]**	**.36 [.16, .57]**	.05 [−.17, .28]	**.53 [.37, .69]**	.11 [−.12, .34]
b_2 (fluid → Y)	.01 [−.24, .26]	.04 [−.25, .32]	.05 [−.26, .37]	.04 [−.19, .27]	.03 [−.30, .35]
b_3 (numeric → Y)	**.34 [.09, .59]**	.27 [−.01, .55]	−.05 [−.36, .26]	**.28 [.06, .50]**	−.04 [−.36, .28]
b_4 (verbal → Y)	.07 [−.20, .33]	.07 [−.22, .37]	−.05 [−.38, .28]	−.04 [−.28, .19]	−.05 [−.39, .29]

Note. Model coefficients are standardized regression weights in the mediation model (see Figure C1); 95% confidence intervals are in brackets and indicate significant effects if an interval does not include zero (marked in boldface); c is the total effect of age on the criterion variable Y (i.e., when potential mediators are ignored); c' is the direct effect of age when mediators are concurrently considered; $a_i \times b_i$ is the indirect effect of age on Y through a specific mediator M_i, while concurrently controlling for all other mediators in the model; indirect-effect estimates are based on bootstrapping (using N = 50,000 samples; Hayes, 2013).

Appendix D

Quality of Individual Model Fits

Goodness of fit of the diffusion model was examined quantitatively and qualitatively. Overall, these evaluations suggested that our selected model variant fit the data from all age groups reasonably well on the individual level and that parameter values could thus be meaningfully interpreted. Goodness of fits was similar to previous applications of the diffusion model in other rapid decision paradigms.

Predicted values of the selected diffusion model variant are plotted against observed values to determine model fit for each participant and across all 8 experimental conditions qualitatively. Panel A shows observed and predicted accuracy (correct acceptances and rejections of gain and loss objects, respectively). Panels B-F show observed and predicted .50, .10, .30, 70, .90 quantiles of correct responses of the RT distributions. Values on the diagonals (with a slope of +1) would indicate perfect fit. As can be seen, the residuals do not indicate systematic deviations in model predictions. Correspondence between observed and predicted values was generally high for both accuracy (all *r*s > .90) and RT (all *r*s > .88 for the selected RT quantiles). Overall, the diffusion model adequately reproduced the effects on individual RTs and on accuracy in the value-based decision paradigm for children as well as younger adults. In each plot, each participant contributed 8 data points across all 4 (Expected Value Level) × 2 (Stimulus Type) experimental conditions.

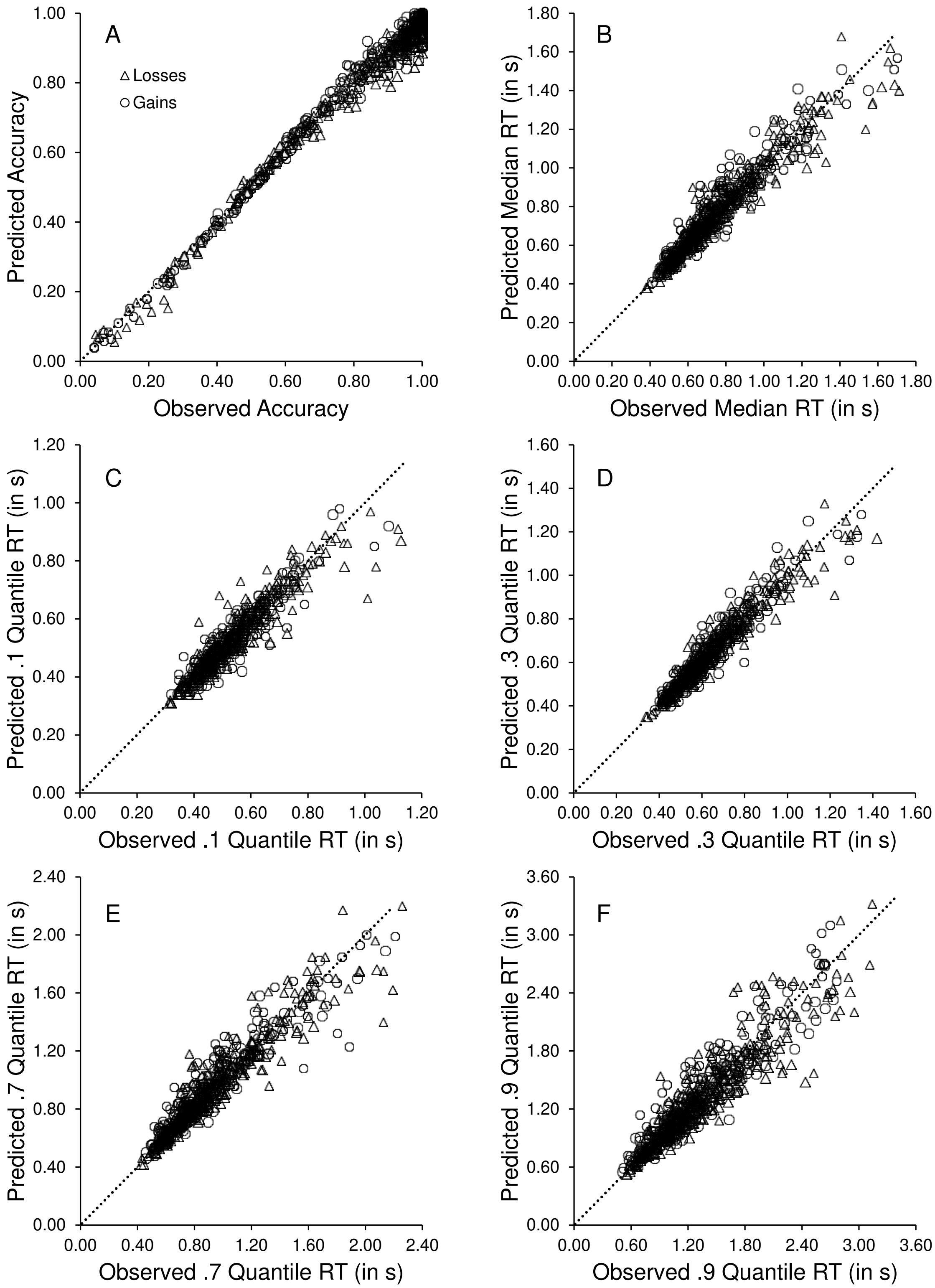


Figure D1. Model predictions and observed accuracy and response-time quantiles.

Supporting Information

Online Supplementary Materials for Article: "Good + Bad = ? Developmental Differences in Balancing Gains and Losses in Decisions from Memory"

Supplement 1. Measures of gain-loss attitudes. Following Tom, Fox, Trepel, and Poldrack (2007), we calculated a behavioral measure of loss aversion (parameter λ) as the ratio of the loss-to-gain impact. Specifically, we performed logistic regression separately for each participant with the magnitudes of the object gain and loss attributes as predictors and accept/reject decisions as the dependent criterion variable. Behavioral loss aversion for each participant was then computed as

$$\lambda = |b|_{\text{loss}} \div |b|_{\text{gain}}$$

where $|b|_{\text{loss}}$ and $|b|_{\text{gain}}$ are the absolute unstandardized individual regression coefficients for the gain and loss attributes, respectively. Parameter λ is similar to the λ measure of the prospect-theory value funtion (Tversky & Kahneman, 1992) with $\lambda > 1$ indicating a relatively stronger impact of losses than gains on decisions. Notably, in the present value-based decision paradigm, median λs were < 1 within all age groups and did not differ between age groups. Figure S1 shows the positively-skewed distributions of individual λs as a function of age group. Figure S2 shows moderate correlations between λ and the diffusion-model bias parameters v_{bias} and z/a.

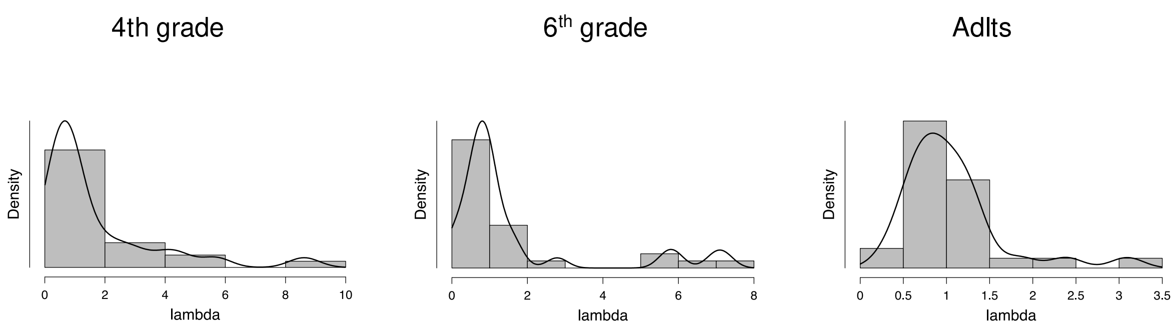


Figure S1. Distributions of individual loss-aversion parameters λ as a function of age group (with λ > 1 indicating loss aversion and λ < 1 indicating gain seeking).

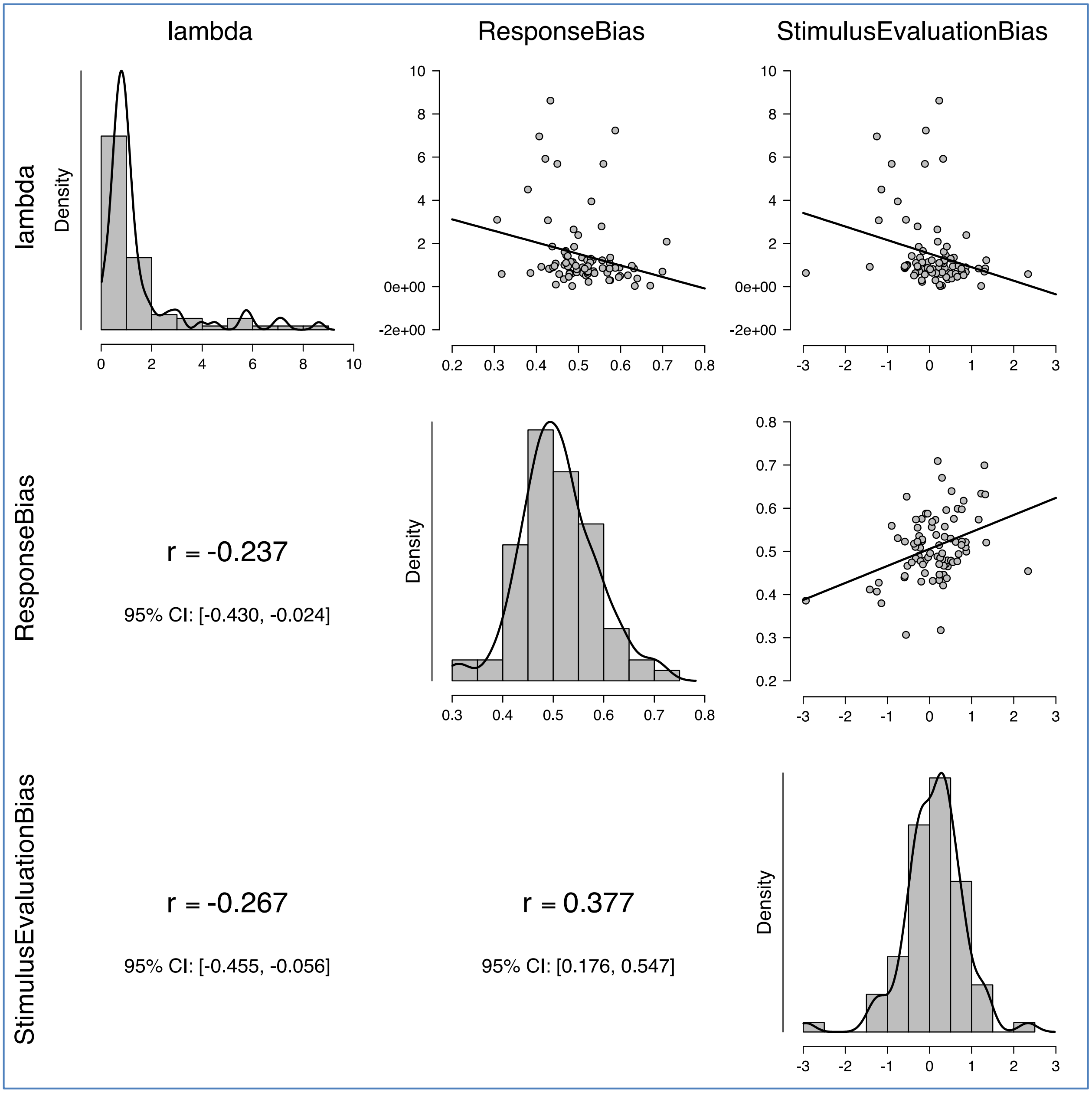


Figure S2. Correlations between individual choice-based measure of loss aversion (parameter λ; cf. Tom et al., 2007) and the diffusion-model estimates of response bias (*z/a*) and stimulus-evaluation drift bias (v_{bias}).

Supplement 2. Zero-order correlations among *z*-transformed tests scales used to construct overarching composite measures. For the mediational analyses (presented in Appendix C of the main article) we computed *z*-transformed unit-weighted composites of average drift, of memory, numerical, and fluid abilities for more stable measurement. As shown below, all constituent scales of a composite measure were correlated (but we also obtained some evidence for a positive manifold; e.g., speed-test scores were not exclusively correlated with other fluid-abilities measures). Figure S3: Relations among fluid-abilities measures. Figure S4: Relations among numerical-abilities measures. Figure S5: Relations among drift rates across domains and value level (graphs created with JASP, 2018).

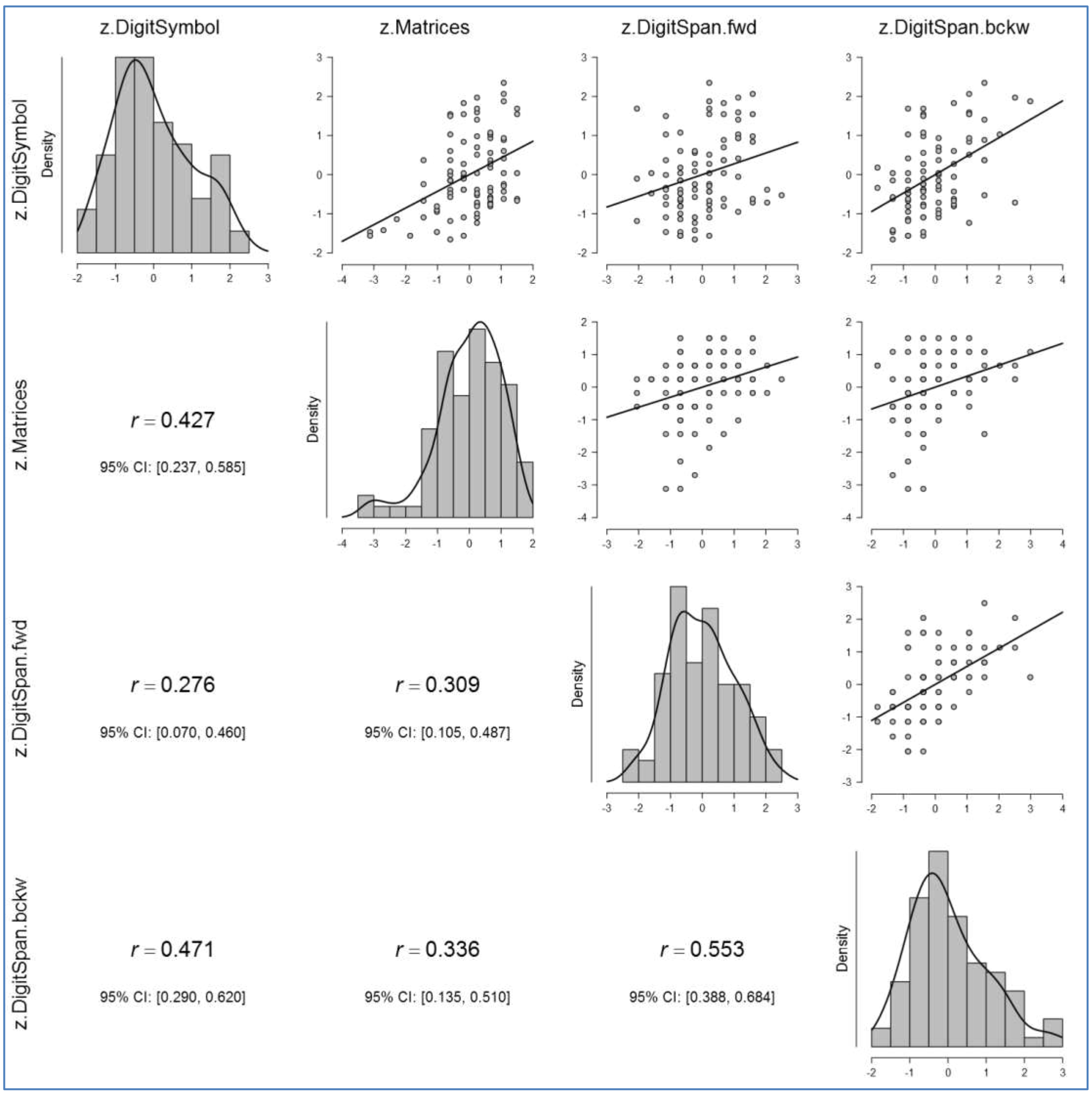


Figure S3. Fluid abilities tests.

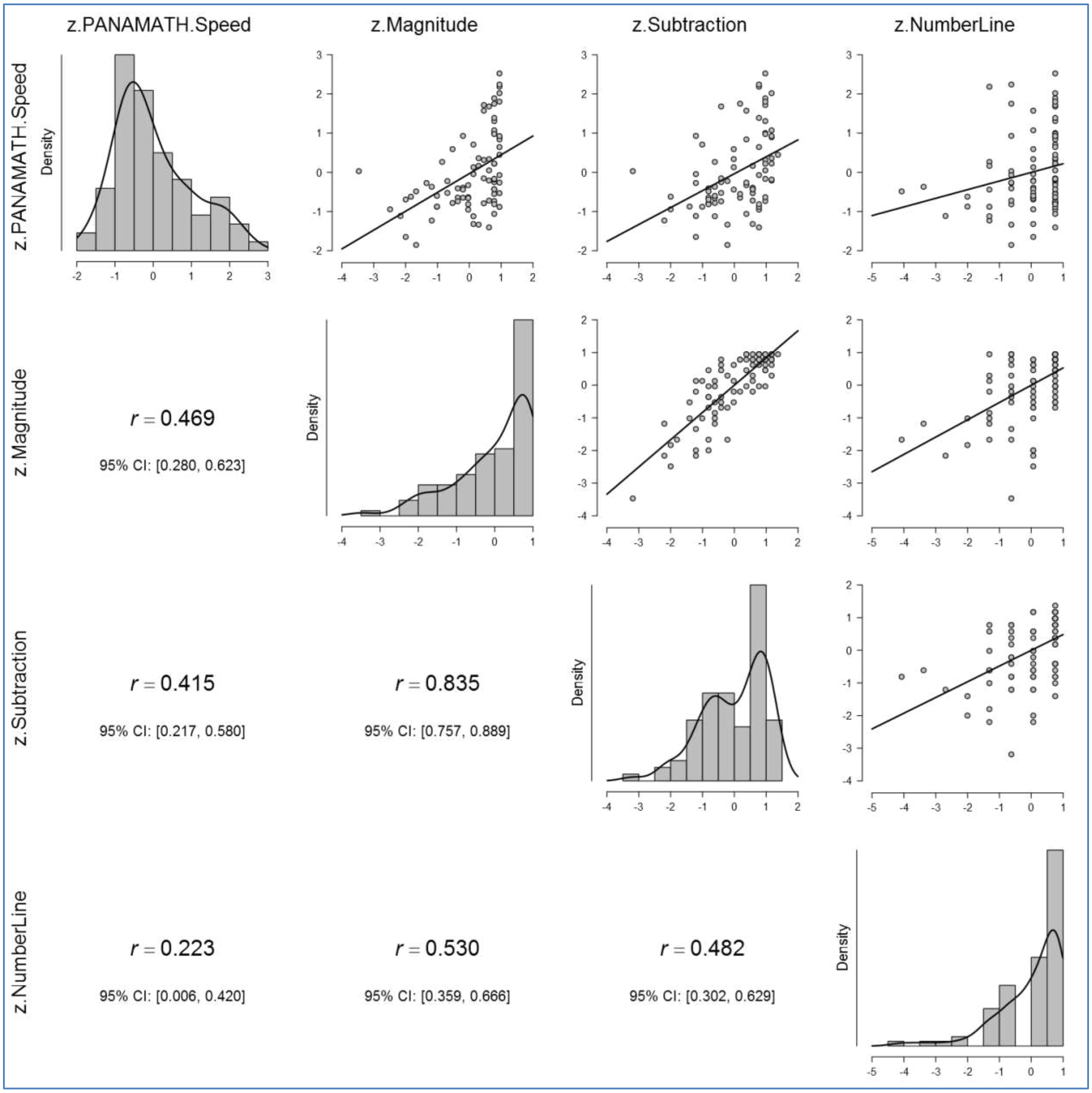


Figure S4. Numerical abilities tests.

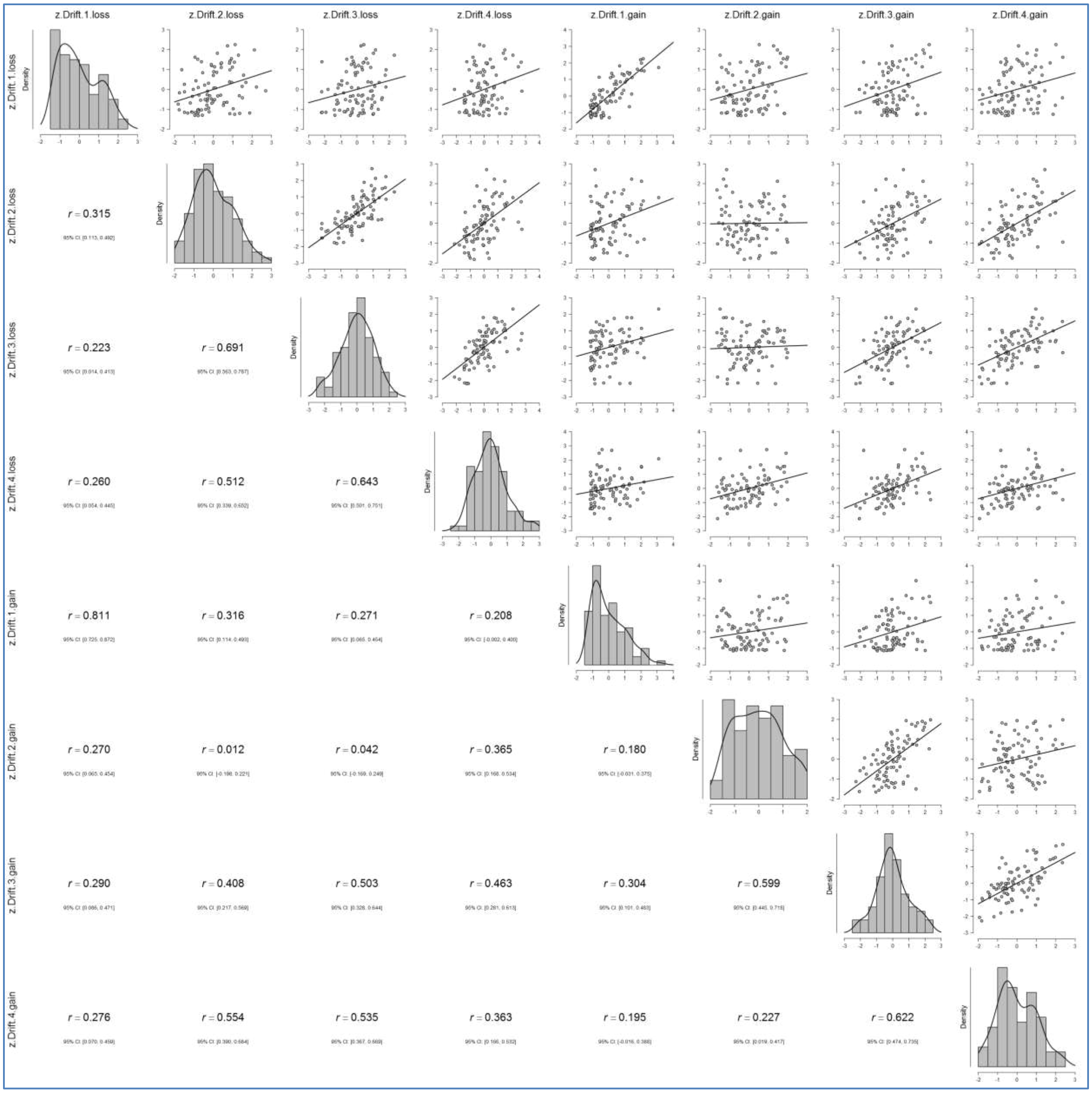


Figure S5. Drift rates.

Supplement 2. Relation between accuracy (proportion of correct decisions), response time (RT) in the gain-loss decision task and diffusion-model parameters drift (*v*), boundary separation (*a*), nondecision time (T_{er}).

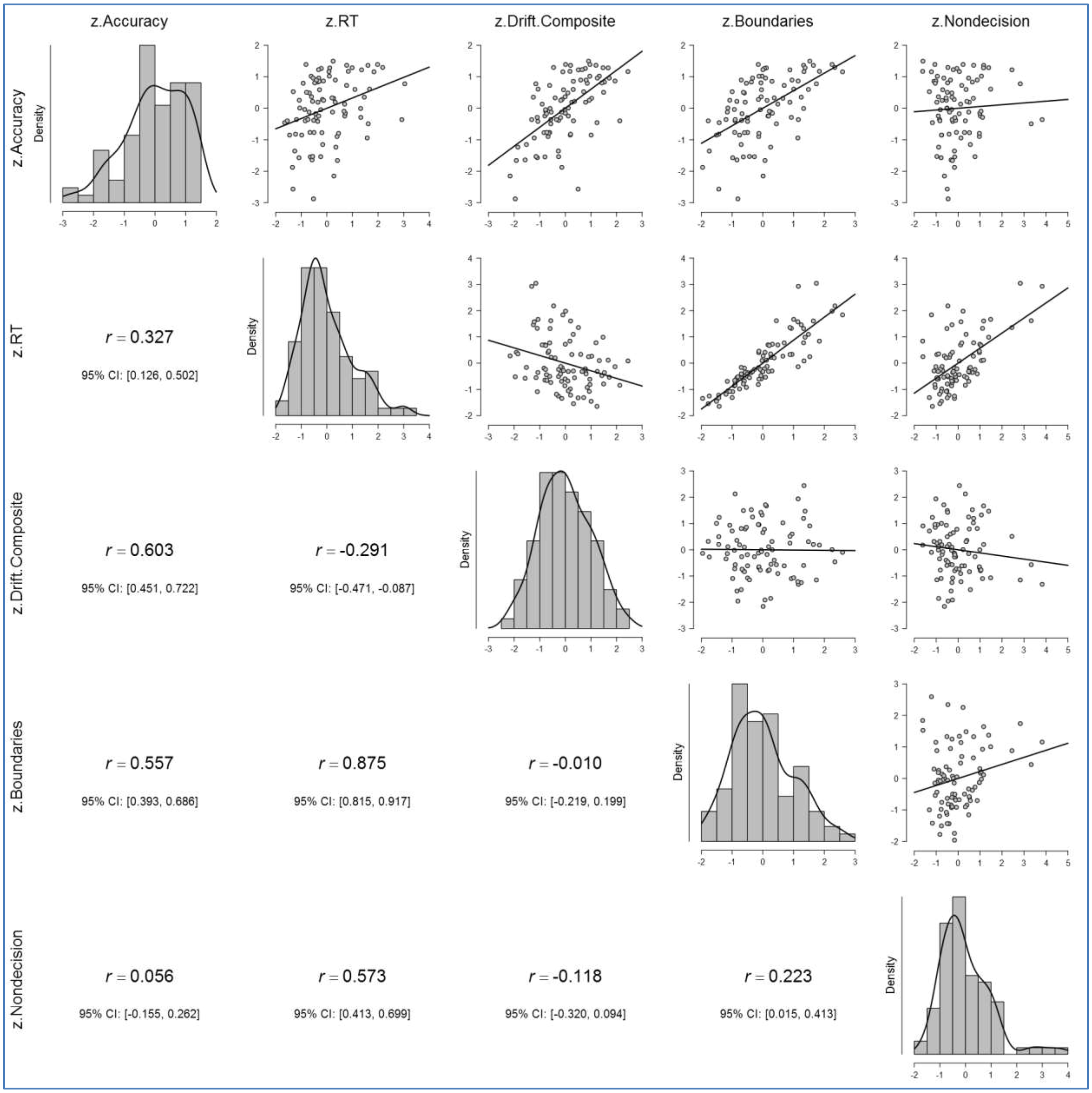
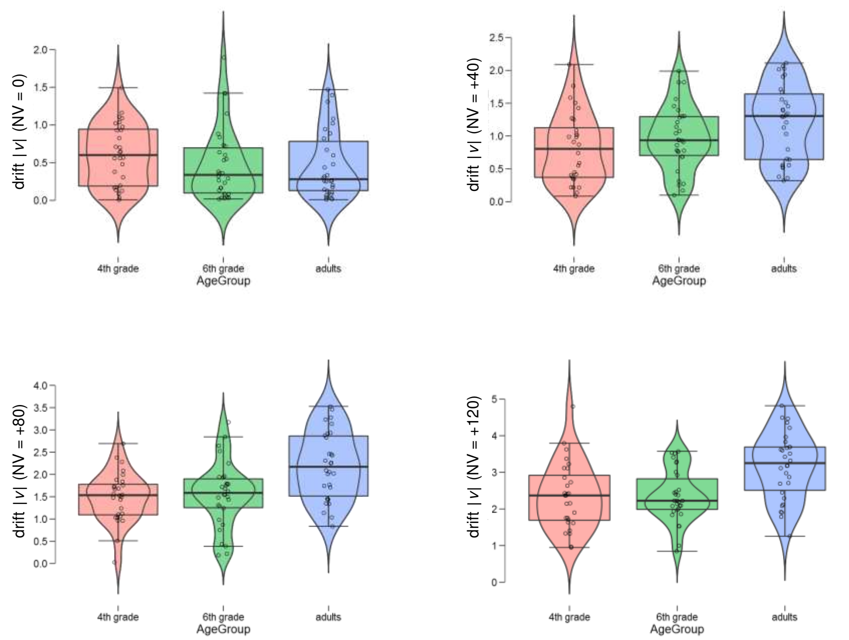


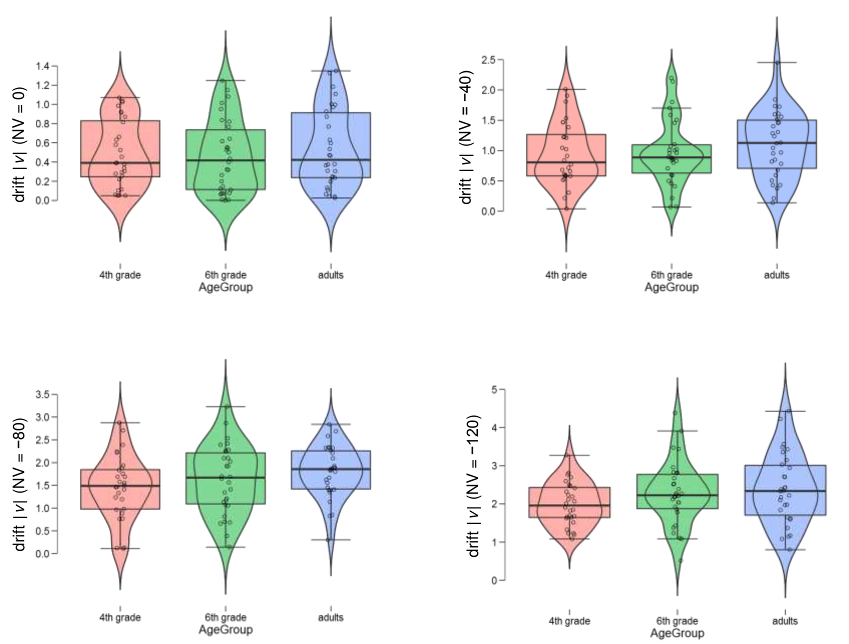
Figure S6. Relation between accuracy, response times, and model parameters.

Supplement 3. Distributions of estimated diffusion-model parameters (combined box-bean plots; circles show the individual parameter estimates) as a function of age group (4th grade, 6th grade, young adults). For drift rates, the absolute values |*v*| are plotted. Consequently, the drift-rate averages in net-value category NV = 0 are slightly larger than zero because positive and negative drifts add up and do not cancel each other out (graphs created with JASP, 2018).

Drift Rate Parameters: Gains



Drift Rate Parameters: Losses



Parameters T_{er}, s_t, a, z/a

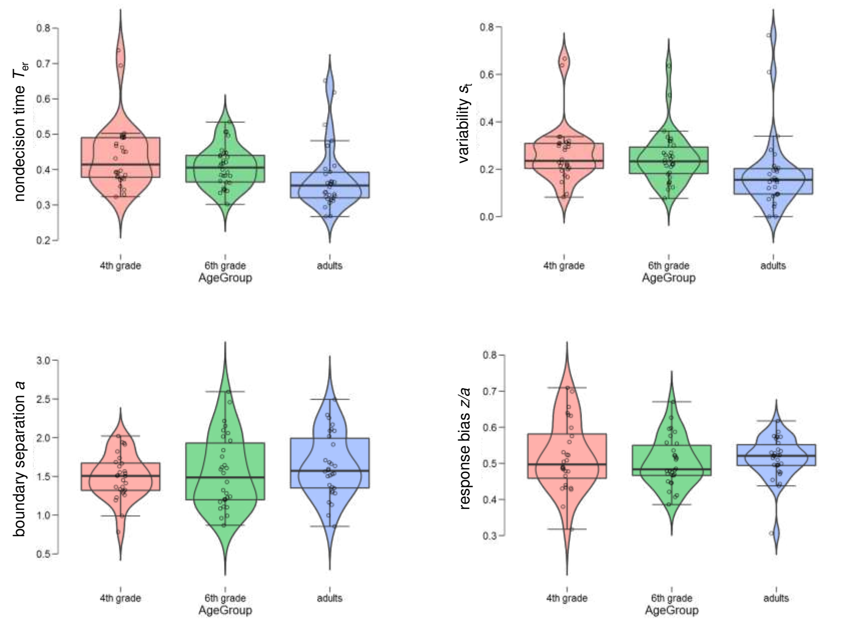


Figure S7. Distribution plots of individual model parameter estimates.

Supplement 4. Description of cognitive test scales and procedures.

To measure fluid abilities, numerical/arithmetic abilities, and vocabulary, all participants completed subscales of standard cognitive tests (including German adaptations of the Wechsler adult intelligence scale WAIS-IV, of the culture fair intelligence test CFT-20, of Panamath, and of arithmetic school tests) that are described in more detail below.

Fluid Abilities Tests

Cognitive Speed

We used a paper-pencil version of the digit-symbol substitution test (adapted from WAIS-IV; Petermann, 2012) as an indicator of cognitive speed. In this speed test, all participants completed as many items on the sheet as possible within 2 min.

Reasoning Matrices

To measure abstract reasoning as an indicator for nonverbal fluid intelligence, participants completed an adapted paper-pencil subscale of the culture fair intelligence test CFT-20 (Weiß, 2006). All participants first completed three examples and then had 3 min to complete up to 15 items of increasing difficulty. The items consisted of 2×2 or 3×3 matrices of geometric patterns with one missing piece in the bottom right corner. On each trial, participants filled in the missing piece by choosing one of five provided patterns.

Digit Span Forward and Backward

Participants heard a list of random numbers, presented by a trained experimenter at the rate of one item per s, and were asked to recall these items in presented order (forward span) or in backward order (backward span). The forward test began with two sequences with a length of two digits and sequence length increased up to nine digits or until two recall errors occurred for a sequence of a given length. The same procedure was then used for the backward test, with a maximum sequence length of eight digits.

Numerical Abilities Tests

Number Line Task

Each participant completed 12 number line estimation problems (e.g., Siegler & Opfer, 2003) that were adapted from a standardized German numerical-abilities test (RZD test; Jacobs & Petermann, 2005). On each trial, a number line (with an upper and lower limit, defined by 2- or 3-digit Arabic numerals) was shown and participants were asked to locate another, symbolically expressed 1- or 2-digit number on the line.

Arithmetic Subtraction Task

In this paper-pencil speed task (adapted from an arithmetic school test, HRT; Haffner, Baro, Parzer, & Resch, 2005), participants completed up to 40 items on a sheet within 2 min. Each item posed a subtraction problem (e.g., “23 – 6 = ____”), where participants had to fill in the blanks with the correct number. Items increased in difficulty (from 1- to 3-digit items).

Arithmetic Comparison Task

In this paper-pencil speed task (adapted from HRT; Haffner et al., 2005), participants completed up to 40 items on a sheet within 2 min. Each item posed a comparison problem (e.g., “2 ____ 817 – 816”), where participants had to fill in the blanks with the correct operator (“=”, “<”, “>”) to make expression a true statement. Items increased in difficulty (from 1- to 4-digit items).

Panamath

The Panamath test was used to explore participants’ “approximate number sense” (Halberda, Ly, Wilmer, Naiman, & Germine, 2012). It has been argued that scores on the Panamath test predict performance on other math achievement tests (e.g., Halberda, Mazzocco, & Feigenson, 2008). We therefore expected number sense to be related to performance in the present arithmetic tasks. In line with this, we found that Panamath performance was significantly correlated with participants’ scores in the arithmetic comparison task ($r = .47$) and arithmetic subtraction task ($r = .42$). Moreover, the correlation between Panamath performance and number-line estimation scores was significant ($r = .22$). All participants worked on a version of the Panamath for 6 min, in which the task difficulty (i.e., four number-ratio levels) was adjusted by the program during testing such that accuracy (but not RTs) was matched across individuals. We therefore used participants’ RTs as dependent variable for further analyses because significant variability between age groups emerged in the RTs (but not in accuracy or the Weber fraction *w*).

Vocabulary

In a paper-pencil vocabulary test, participants completed 30 items of increasing difficulty. Each item consisted of one probe word (e.g., “scarf”) and participants were asked to choose a semantically matching target out of five other words with the closest meaning connection (e.g., “tie”).

References Online Supplement

Haffner, J., Baro, K., Parzer, P., & Resch, F. (2005). *Heidelberger Rechentest HRT 1-4*. Göttingen, Germany: Hogrefe.

Halberda, J., Ly, R., Wilmer, J. B., Naiman, D. Q., & Germine, L. (2012). Number sense across the lifespan as revealed by a massive Internet-based sample. *Proceedings of the National Academy of Sciences, 109,* 11116–11120.

Halberda, J., Mazzocco, M. M. M. & Feigenson, L. (2008). Individual differences in non-verbal number acuity correlate with maths achievement. *Nature 455,* 665–668.

Jacobs, C., & Petermann, F. (2005). *Rechenfertigkeiten-und Zahlenverarbeitungs-Diagnostikum für die 2. bis 6. Klasse: RZD 2-6.* Göttingen, Germany: Hogrefe.

JASP Team (2018). *JASP* (Version 0.9) [Computer software].

Petermann, F. (2012). *Wechsler Adult Intelligence Scale* WAIS-IV (deutsche Version). [German Version of the Wechsler Adult Intelligence Scale]. Frankfurt, Germany: Pearson.

Siegler, R.S., & Opfer, J. (2003). The development of numerical estimation: evidence for multiple representations of numerical quantity. *Psychological Science, 14,* 237–243.

Tom, S. M., Fox, C. R., Trepel, C., & Poldrack, R. A. (2007). The neural basis of loss aversion in decision-making under risk. *Science, 315,* 515–518.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.

Weiß, R. H. (2006). *CFT 20-R: Grundintelligenztest Skala*. [Culture Fair Intelligence Test 20] Göttingen, Germany: Hogrefe.